AD-A049 955   PATTERN ANALYSIS AND RECOGNITION CORP ROME N Y   F/G 15/4
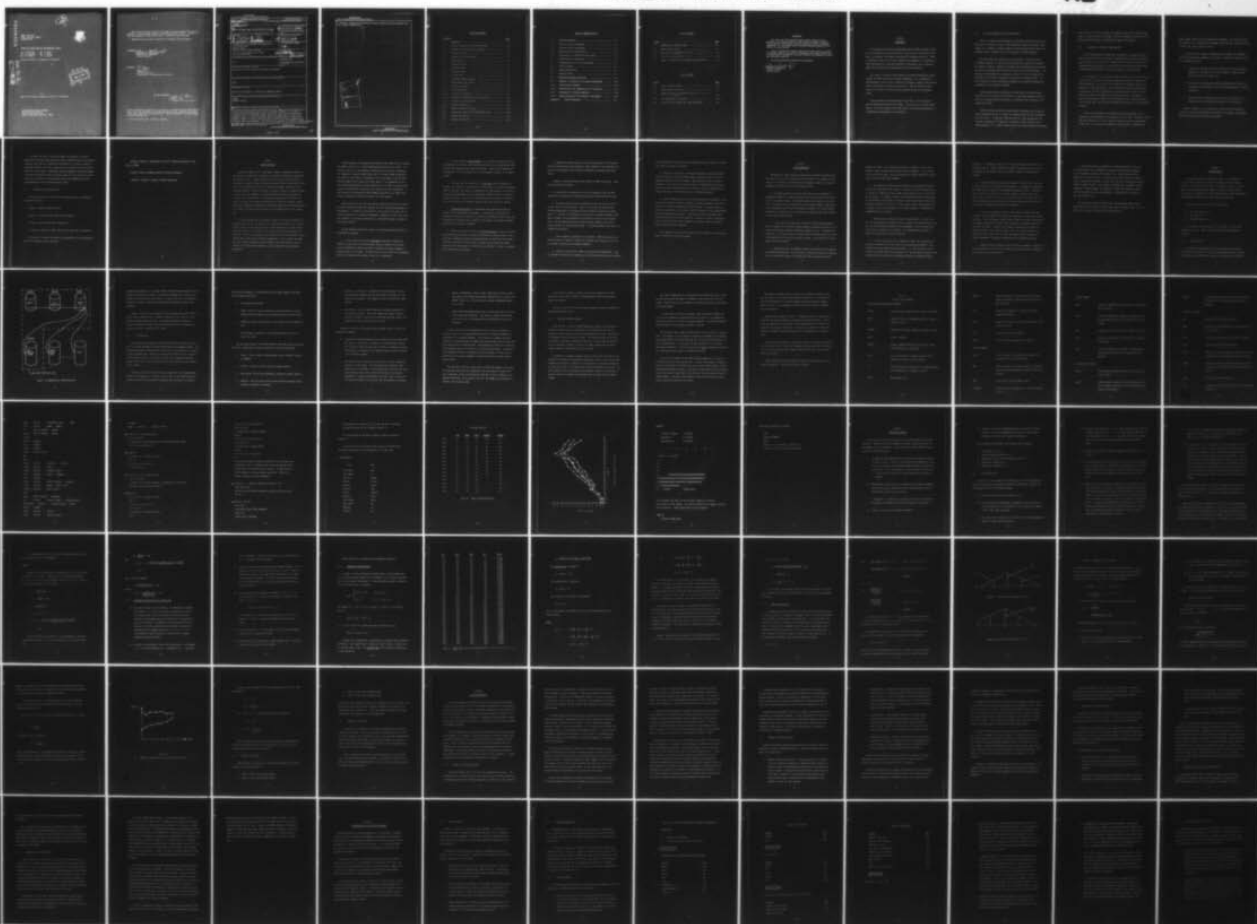   TREND AND ERROR ANALYSIS METHODOLOGY SYSTEM.(U)
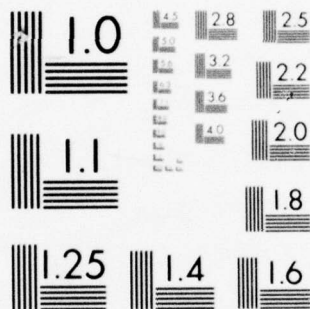   DEC 77   D BIRNBAUM, B BUDELMAN, R D'AMORE   F30602-76-C-0206
UNCLASSIFIED   RADC-TR-77-416
                                                                                 NL

1 OF 2
AD
A049955

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

RADC-TR-77-416
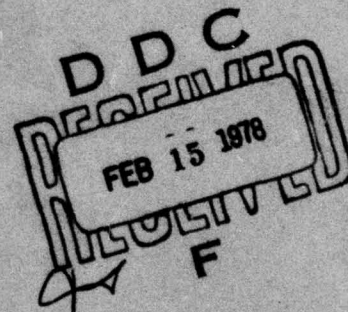Final Technical Report
December 1977

TREND AND ERROR ANALYSIS METHODOLOGY SYSTEM

Dr. D. Birnbaum   Mr. S. Haehn
Mrs. B. Budelman   Dr. J. Morris
Mr. R. D'Amore   Mr. J. Wilson

Pattern Analysis & Recognition Corporation

DDC
FEB 15 1978
F

Approved for public release; distribution unlimited.

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, New York 13441

This report has been reviewed by the RADC Information Office (OI) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-77-416 has been reviewed and is approved for publication.

APPROVED: *Patricia M Langendorf*

PATRICIA M. LANGENDORF
Project Engineer

APPROVED: *H Davis*

HOWARD DAVIS
Technical Director
Intelligence & Reconnaissance Division

FOR THE COMMANDER: *John P Huss*

JOHN P. HUSS
Acting Chief, Plans Office

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (IRDA) Griffiss AFB NY 13441. This will assist us in maintaining a current mailing list.

Do not return this copy. Retain or destroy.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER RADC-TR-77-416	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TREND AND ERROR ANALYSIS METHODOLOGY SYSTEM		5. TYPE OF REPORT & PERIOD COVERED Final Technical Report
		6. PERFORMING ORG. REPORT NUMBER N/A
7. AUTHOR(s) D. Birnbaum S. Haehn B. Budelman J. Morris R. D'Amore		8. CONTRACT OR GRANT NUMBER(s) F30602-76-C-0206
9. PERFORMING ORGANIZATION NAME AND ADDRESS Pattern Analysis & Recognition Corporation 228 Liberty Plaza Rome NY 13440		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 96820053
11. CONTROLLING OFFICE NAME AND ADDRESS Rome Air Development Center (IRDA) Griffiss AFB NY 13441		12. REPORT DATE December 1977
		13. NUMBER OF PAGES 94 103p.
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Same		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Same

18. SUPPLEMENTARY NOTES

RADC Project Engineer: Patricia M. Langendorf (IRDA)

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

TEAMS
DIPPOLS system

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Trend and Error Analysis Methodology System (TEAMS) has been developed for the Defense Intelligence Directorate of Estimates (DIA-DE-1) to provide a computer-based data management system which will (a) store past projections prepared by DE-1, (b) perform simple statistical operations to detect and quantify errors in these projections, and (c) provide output in tabular and graphic formats for use in internal quality control and in the preparation of briefing charts and diagrams. The final report provides an overview of TEAMS. → over

DD FORM
1 JAN 73 1473   EDITION OF 1 NOV 65 IS OBSOLETE

In addition, Program Specifications and User's Manual have been prepared for more limited distribution.

ACCESSION for

NTIS	B ti Section
DDC	B ti Section
UNANNOUNC D	
JUSTIFICATION	

BY
DISTRIBUTION/AVAILABILITY CODES

| Dist. | CIAL |

A

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

LIST OF FIGURES

LIST OF TABLES

EVALUATION

The Trend and Error Analysis System (TEAMS) defined by this contract is an effective tool to enable people performing estimates for the Defense Intelligence Plans and Production (DIPP) to evaluate themselves. It will enable them to publish objective measures of the quality of their product.

This is particularly important because the DIPP itself is becoming the primary planning guide both in determining the threat to which the US forces must be responsive, but also to define the opponents in gaming and simulation exercises.

No previous objective measures were available.

Patricia Langendorf
PATRICIA LANGENDORF
Project Engineer
Intel Sys Sec

SECTION 1

BACKGROUND

The Trend and Error Analysis Methodology System (TEAMS) provides a facility for the detection and quantification of past errors in projections developed by the Defense Intelligence Directorate of Estimates (DIA-DE-1). TEAMS is intended for the use of DIA-DE-1 estimators and managers in reviewing and improving methods for long-range estimative intelligence, and in producing graphs and briefing charts documenting DE-1's past performance.

The scope of the effort reported here has included comprehensive system designs for TEAMS, but has not included programming. It is intended for implementation on DIA's HIS-635 computer at Arlington, Virginia, under the HIS General Comprehensive Operating System (GCOS). Work was performed under Contract No. F30602-76-C-0206, and was monitored by Rome Air Development Center.

Three documents are being produced at this time. One of these will provide Program Specifications for TEAMS. The second will be a draft User's Manual. The third document, this Final Report, will provide background and supplementary documentation for the project.

1-1

1.1. THE DIPP VOLUMES AND THE DIPPOLS SYSTEM

DIA-DE-1 is tasked with development of long-range projections for DOD and other users. Working with many, often conflicting, sources of finished intelligence, DE-1 produces a unified estimate of future foreign force levels; such estimates play an essential role in this country's military planning.

Four major areas are covered in DE-1 projections: the Soviet Union, the People's Republic of China, Non-Soviet Warsaw Pact nations, and Selected Non-Communist nations. Detailed projections for the present year and the next ten years are produced; these consist of a High, Low, and Best estimate. (The meaning of these estimates is discussed in Section 7.) DE-1 projections are published as the Defense Intelligence Projections for Planning (DIPP). The DIPP volumes also include substantial background and supporting information, in addition to the numerical estimates.

Selected quantitative information and annotations are made available through the DIPP On-Line Service (DIPPOLS), a large data management system in which DE-1 provides updated projections to supplement those contained in the annual DIPP volumes.

DIPPOLS provides facilities for data update and access, and through its report generating routine is capable of producing tables for use in preparing the DIPP volumes. In addition, some facilities for data manipulation are included, permitting the combination of specified weapon systems, use of scaling factors, etc. DIPPOLS stores the most up-to-date estimates of military

force levels for the next ten years, the present year, and the preceding four years; as these estimates are revised, the new information is entered into the data base, and the older, now obsolete, estimates are discarded.

1.2.   THE NEED FOR TREND AND ERROR ANALYSIS

It has recently become apparent that some method is required for evaluating the quality of DE-1 projections. Congressional and military groups have been concerned for the accuracy of information which plays a crucial role in long-range military planning. In addition, DE-1 requires an effective means of determining the sources of errors which appear in its projections.

This information is not readily available, however, since records are not stored in such a way as to permit easy comparison of earlier estimates with later information. As a result, DE-1 estimators and managers have found it necessary to spend substantial time in searching through historical records, extracting the required information, and preparing tables, charts, and diagrams to be used for briefings and for internal reviews. The amount of clerical work needed to produce this information has had a serious impact on DE-1's primary task, the continued development and production of intelligence estimates.

Trend and error analysis was nevertheless seen as essential to the preservation and enhancement of the quality of DE-1's work. The goal of trend and error analysis is to locate potential sources of error in intelligence projections. In addition, the tendency to overestimate or underestimate

weapon system levels should be detected and eliminated. By providing contin-
uous feedback to the estimators and managers, trend and error analysis provides
a vehicle for quality control in DE-1.

The Trend and Error Analysis Methodology System (TEAMS) was designed to
provide computer-based assistance to this task. Specifically, TEAMS provides
the following facilities:

1. A historical data base which will contain DE-1's past projections,
    together with estimates of the actual numbers of weapon systems
    derived from orders of battle (OB).

2. Facilities for combining and comparing estimates and OB data, for
    locating and scoring errors, and for developing measures of bias and
    uncertainty.

3. Output routines to produce tables and charts as required for DE-1
    analyses, and for briefings and reviews of DE-1 performance.

TEAMS is intended to be an interactive, easily-used data management
system, for use by non-statistically oriented DE-1 estimators and managers;
both on-line and batch facilities are provided.

1.3.   THE ROLE OF TEAMS

TEAMS will assist DE-1 in evaluating its own performance. In addition,
TEAMS will reduce the time required for preparation of charts and tables used
in briefings.

Through the use of TEAMS, DE-1 will have access to data concerning its
past performance, which will provide essential feedback to the estimation
process. When errors are located, they can be further reviewed to determine
whether they represent a continuing trend, or whether they reflect an anoma-
lous situation due to non-recurring historical factors.

A central concept in the development of TEAMS has been the measurement of
error. This concept is discussed in more detail in Section 7 of this report,
in which DE-1 methodology is reviewed.

Briefly, it should be emphasized that the concept of error in long-range
intelligence projections does not generally represent a failure in the estima-
tive process. An estimate may be entirely correct, in terms of the intentions
and capabilities of a potential adversary. Nevertheless, these intentions and
capabilities may change in the future, in ways which are unpredictable in
principle. For this reason, an "error" may simply represent a change in the
intentions of a potential adversary, which could not have been foreseen at
the time the projection was made.

It should also be noted that the scoring rules provided with TEAMS have a particular function in providing assistance to DE-1 estimators and managers. It would be seriously counter-productive to use these scores for rating the quality of the work performed by the individual analyst. It is inevitable, given the great variety of weapon systems and nations for which DE-1 must produce estimates, that some tasks will be considerably more difficult than others. Data for some systems are extremely scarce and of doubtful quality. Under these circumstances, the more difficult estimates are likely to receive lower scores than the easier estimates. If scores for these estimates were then interpreted as suggesting that the work performed in producing the estimates was of poorer quality, this would place an unwarranted stigma upon those estimators who had taken on the most difficult tasks. Instead, the scores are intended to assist in identifying problem areas in the DE-1 projections.

The development of a scoring rule has played an important role in the TEAMS project. In earlier attempts to evaluate DE-1 projections, a "hit or miss" system of scoring was used, which consisted simply of a count of the number of times projections had been correct, compared with the number of times they had been wrong. If this method of scoring were used over a long period of time, it would have an unfortunate effect, since an estimator can improve such a score simply by increasing the "spread," or the distance between the High and Low estimates that he produces. By making his projections less precise, he has a greater chance of being correct. Unfortunately, of course, his projections would then contain less information for the ultimate user of the intelligence, and would be of less value.

A scoring rule should instead encourage the estimator to produce a spread which accurately represents the range of possible values which might be expected, given the best information now available. As noted in Section 7, estimators attempt to provide a range which will enclose the true values in three out of four cases. Estimations should, therefore, receive the highest score when the range includes the correct score in 75 per cent of the cases. With such a scoring rule, the estimator will not be rewarded either for underestimating or for overestimating the range.

1.4.   OVERVIEW OF THE FINAL REPORT

This final report will include the following sections, following this introductory section:

Section 2 presents TEAMS objectives.

Section 3 will describe the TEAMS design strategy.

Section 4 will summarize TEAMS capabilities.

In Section 5, samples of TEAMS sessions with users will be presented.

In Section 6, a further discussion and presentation of the Scoring Rule and other statistics will be included.

Section 7 contains a description of DIA-DE-1 projection methods as they apply to TEAMS.

Finally, Section 8 suggests areas for future development.

Appendix A contains a summary of TEAMS terminology.

SECTION 2

TEAMS OBJECTIVES

During the summer of 1974 there began a period of increased interest in the accuracy of information contained in the Defense Intelligence Projections for Planning (DIPP). These projections serve as a basis for DOD policy by providing a consensus judgment concerning foreign military force levels at the time of the projections and for the following ten years. Because of this increased interest, DIA-DE-1 was asked to produce a variety of summary tables, briefing charts and diagrams, and other materials required to document the accuracy of past projections. A considerable expenditure of effort was required to extract the necessary information from historical records, perform appropriate mathematical operations, and produce reports or briefing materials. Such efforts have had an obvious impact on the primary tasks of DE-1 personnel.

It became apparent that an automated system for storage and retrieval of required data could be cost-effective in reducing the personnel time involved in the analysis of past projections and in the production of charts, tables and other outputs. In addition, such a system could provide information on the basis of which DE-1 personnel could evaluate their own performance, locate problem areas, and improve projection methodology. Such a system, the Trend and Error Analysis Methodology System (TEAMS), was envisioned as a tool to assist DE-1 estimators and managers in report generation and performance evaluation.

Pattern Analysis and Recognition Corporation (PAR) began work on a design

for TEAMS in April, 1976. Three phases were projected for the effort: (a)

initial study of DE-1 requirements, methodology, and personnel backgrounds;

(b) analysis of required capabilities; and (c) system design development.

The result of this effort is contained in this Final Report, with the accom-

panying Program Specifications and User's Manual. No programming for the

final TEAMS implementation has been performed, but a number of programs were

written to test specific algorithms, data formats, and user interaction. A

small system for demonstrating TEAMS commands was written, the TEAMS Pilot,

which is operational under GCOS on RADC's HIS 6080 computer.

Design objectives for TEAMS have evolved during the course of the effort,

as project personnel became more familiar with DE-1 operating procedures and

requirements. In general, the direction of this evolution has been that of

providing designs for an effective management information system, which would

have efficient file-access procedures and a comprehensive range of graphical

and tabular outputs.

As here designed, TEAMS will provide the following capabilities for DE-1

estimators and managers:

a. It will maintain an on-line <u>data base</u> consisting of historical

records of projections for foreign weapon systems, together with the most

accurate information currently available concerning the actual numbers of

weapon systems in past years. The data base will thus contain all information

required to determine the accuracy of past DE-1 projections.

b. It will provide a <u>user interface</u>, i.e. a group of programs that will ask questions of the user, accept answers from him, call the appropriate TEAMS routines for processing, and output the results. Output can be directed to a user terminal, to a line printer, or to an incremental plotter, at the user's request.

c. Each user will be provided with a <u>work area</u>, which functions as a notebook for storage of information from the data base, additional information added by the user, the results of computations or modifications of the data by the user, and information which is temporarily stored for output to the line printer or plotter. Any or all of the work area may be erased or modified by the user at any time; it otherwise remains intact from session to session.

d. <u>Statistical functions</u> for identifying and quantifying trends and errors in DE-1 projections are provided. A scoring rule is available to estimate and compare the quality of projections; estimates of error, bias, and uncertainty are provided; and a measure of the degree of correlation between any two sets of scores is included.

e. TEAMS will provide a variety of <u>charts and graphs</u>, which will compare DE-1 projections with order of battle (OB) estimates, estimates made at various times, estimates of groups of weapon systems, and other data of interest. The user is provided with the ability to design his own graphs and plots, using any data that appear in the data base; in addition, a number of standard plot formats will be provided.

f. TEAMS also provides outputs to the line printer and to the terminal. These can include tables and listings of data contained in the data base, the results of statistical tests, and other information concerning past projections.

A number of general principles have guided the TEAMS development. These have included the following:

a. Users would be provided with tools for evaluating their own work. They should be assisted in locating and correcting various sources of error.

b. The user interface should be adjusted to meet the needs of users at various levels of sophistication. Users of the system are provided with complete explanations of available functions simply by typing in a question mark, "?". Because of the relatively slow output of terminal equipment available for TEAMS, such explanations are provided only when the user requests them. The more experienced user can enter several commands on one line, without waiting for prompting from TEAMS. To minimize typing, brief forms for commands are available.

c. Human judgment is enhanced, not superseded. TEAMS provides the user with facilities for testing a variety of hypotheses and hunches, but it does not provide a substitute for informed judgments.

d. Statistics included with TEAMS are simple and straightforward. They are intended to provide DE-1 estimators, few of whom are statistically-oriented,

with measurements of error, bias, and uncertainty which can easily be checked with the aid of a small calculator.

e. Within the limitations of available equipment, outputs from TEAMS should be usable as briefing charts, graphs, and tables, with little or no additional art work or hand lettering. When existing plotter or printer equipment is not capable of producing outputs of sufficiently high quality, the outputs should nevertheless serve as guides to artists, draftsmen, etc., without substantial further editing by DE-1 technical personnel.

f. Modern programming practices, including structured programming, have been shown to provide substantial savings in program development time and in the costs of software maintenance and modification. A top-down design approach has been used in this project, which is consistent with recommendations for modern programming practices. Specifically, system requirements and system specifications have been developed in advance of coding, eliminating much of the wasted motion that frequently occurs when coding is initiated before designs have been completed.

These general objectives were combined with an approach to system design which is described in the next section.

SECTION 3

DESIGN STRATEGIES

Development of TEAMS designs has required the solution to several problems which will be briefly summarized in this section. The purpose of this section is to provide an introduction for the general reader to the type of strategies that have guided the TEAMS design effort.

a. The TEAMS design should provide the user with the capability of expanding the system to include additional functions which may not have been part of the initial design. The EXECUTE command was included for this purpose. Through the use of EXECUTE, the user can request TEAMS to perform any set of operations that the user has previously defined. He is not limited, therefore, to a restricted set of routines that happen to have been included in the initial system designs.

b. It is likely that TEAMS users will want to input textual material which is longer than the rather limited number of characters available on one line of the terminal equipment currently available to DE-1. To provide this capability, it was made possible to input lines of arbitrary length, such that both commands and data could be stored as strings. User comments can also be stored in the same format.

c. TEAMS should also be expansible through the addition of new programs and enhanced capabilities. For this reason, programs are modular in organization, to minimize the amount of revision that will be required when new

segments are added. The statistical routines, for example, are all called

through a single program employing user-supplied parameters. As new statis-

tical functions are added, only this program will require revision to provide

calls to the new statistical programs.

    d. More generally, TEAMS has been designed on the assumption that DE-1

policies and requirements were subject to change, and that the system should

be capable of changing with them, rather than locking DE-1 into a restricted

set of options. For example, there is some possibility that DE-1 projections

may be extended in the future to include 25 years, rather than the ten years

presently contained in the DIPP. If this change is made, TEAMS programs may

be adapted through the modification of a single parameter; no extensive re-

programming will be required.

    e. TEAMS has been designed for ease in implementation. For this pur-

pose, an instrumentation package has been included in the system design, which

will provide information required by the programmer to locate coding errors

and to test all portions of the system. The instrumentation package can

easily be disabled when testing is complete.

    f. The user can not only issue commands to TEAMS, but also obtain in-

formation concerning the current state of the operating environment. Such

commands as the following have been included: (1) MENU provides a list of

currently available options in the system; (2) MAP will give a picture of the

current state of the work area, indicating the amount of additional space

available; (3) VARIABLES returns a list of the variables being stored in the work area; and (4) LENGTH indicates the length of variables, both in a form appropriate to the variable type and in terms of amount of storage space required.

g. Frequent integrity checks are performed on the work area. One type of check is the use of redundant code in pointers. A pointer is contained in the symbol table, pointing to each variable represented there; in addition, a second pointer is contained in the stored variable, pointing back to the symbol table. The sole purpose of this second pointer is to insure that the correct variable has been located, and to prevent errors that might occur if the variable had somehow been erased or mislocated.

h. To prevent accidental erasure of data, several commands have been included. LOCK and INHIBIT will prevent the destruction of any data to which they apply. When the general command CLEAR is given, the system requires that the command be repeated, to reduce the possibility of accidental erasure of the entire work area. Finally, the "erasure" of data does not destroy it; only the pointers to the data are reset. Thus, if a user accidentally executes the CLEAR command, the work area can nevertheless be restored by a system programmer if subsequent operations have not destroyed the data.

i. TEAMS provides a flexible syntax for control statements. Examples of this flexibility are found in the pattern string used in the PLOT command and the control string used in MODIFY.

j. The system should be insensitive to peripheral protocols; that is, the system should make it possible to substitute new peripheral equipment, such as terminals, plotters, CRT displays, etc., without extensive reprogramming of the system itself. This goal is achieved by handling all input/ output to peripheral equipment, the work area, and batch processing facilities through a single program. Thus, changes in hardware can be effected through changes in a single program. For example, if TEAMS were modified to permit use of both upper and lower case characters, a change would be required in only one routine, TIO.

The first three sections of this report have summarized general TEAMS objectives and the strategies that have been used to implement them. The next section will present an overview of the system itself.

SECTION 4

SYSTEM OVERVIEW

This section provides an overview of TEAMS. It is intended to
give the reader a general picture of the system, without an excess
of detail. Potential users of TEAMS should obtain a copy of the User's
Manual, which describes system commands in detail and gives samples
of their use. Programmers and others responsible for TEAMS implementation
and maintenance should refer to the Program Specifications.

This section includes the following subsections:

a. Data Base Design (4.1.)

b. Work Area (4.2.)

c. Command Language (4.3.)

The following section, Section 5, will present a sample TEAMS
session with a user. TEAMS statistical functions are described in
Section 6.

4.1.   DATA BASE DESIGN

The TEAMS data base design, which has evolved over the term of
the project, is intended to provide for flexible and efficient storage
of various data, including both descriptive information and numerical

data. The TEAMS data base (TDB) is primarily intended to provide sufficient information to support the various analysis/display functions currently planned. However, the design is not necessarily restricted to meeting only current needs. Evolutionary growth in TEAMS is supported by allowing additional data to be stored within the existing structures.

The TEAMS data base uses information derived from DIPPOLS (and extracted by hand from earlier DIPP volumes). This consists of High, Low, and Best estimates made during a given projection year for that year and for the ensuing ten years. In addition, it contains OB data for each of the preceding four years. Information in the OB is based on estimates of the numbers of foreign weapon systems. As is true in all areas of intelligence, such estimates are subject to error. It is assumed, however, that the most recent OB for any given year represents the best information that is obtainable, and so is called the "Best OB." This is used in determining the amount and direction of errors in earlier projections.

The TDB consists of four random files:

a. DIPPLIST, which contains a list of the DIPP's currently available, with pointers to the TABFILE.

b. TABFILE, which contains lists of the Table ID's (weapon systems), with pointers to DATREC.

c.   DATREC, which contains Record ID's together with pointers to the various data records.

d.   DATFILE, which includes the actual data, consisting of historical projections from current and past DIPP's.

As shown in Figure 4-1, the data base can be expanded simply by including additional pointers in DATREC, which will designate the locations on other disk files which contain the additional data.

TEAMS must provide for storage of both historical and current data. This normally results in data records which are chronologically stored (by projection year) in the data file. In addition, directory information (in the pointer files) keeps track of the data storage status as well as the data categories.

Data may be incomplete. For example, during certain years no projection at all may have been made. In another case, the projection was made but has not yet been included in the TEAMS data base. In other cases, the OB information may not yet be available for certain years. The TDB design, together with the data retrieval routines (primarily SELECT) are structured in such a way as to keep track of and account for these irregularities in data storage.

In current designs, projection data for various weapon systems (Table ID's) are stored according to DIPP. TEAMS provides for another level of data

DIPPLIST   TABFILE   DATREC

DIPP's

CC's
TABLE ID's

RECORD
ID's

DATFILE

DATA
(PROJEC-
TIONS)

OTHER
TEAMS
DATA

DATA

CURRENT TEAMS DATA BASE
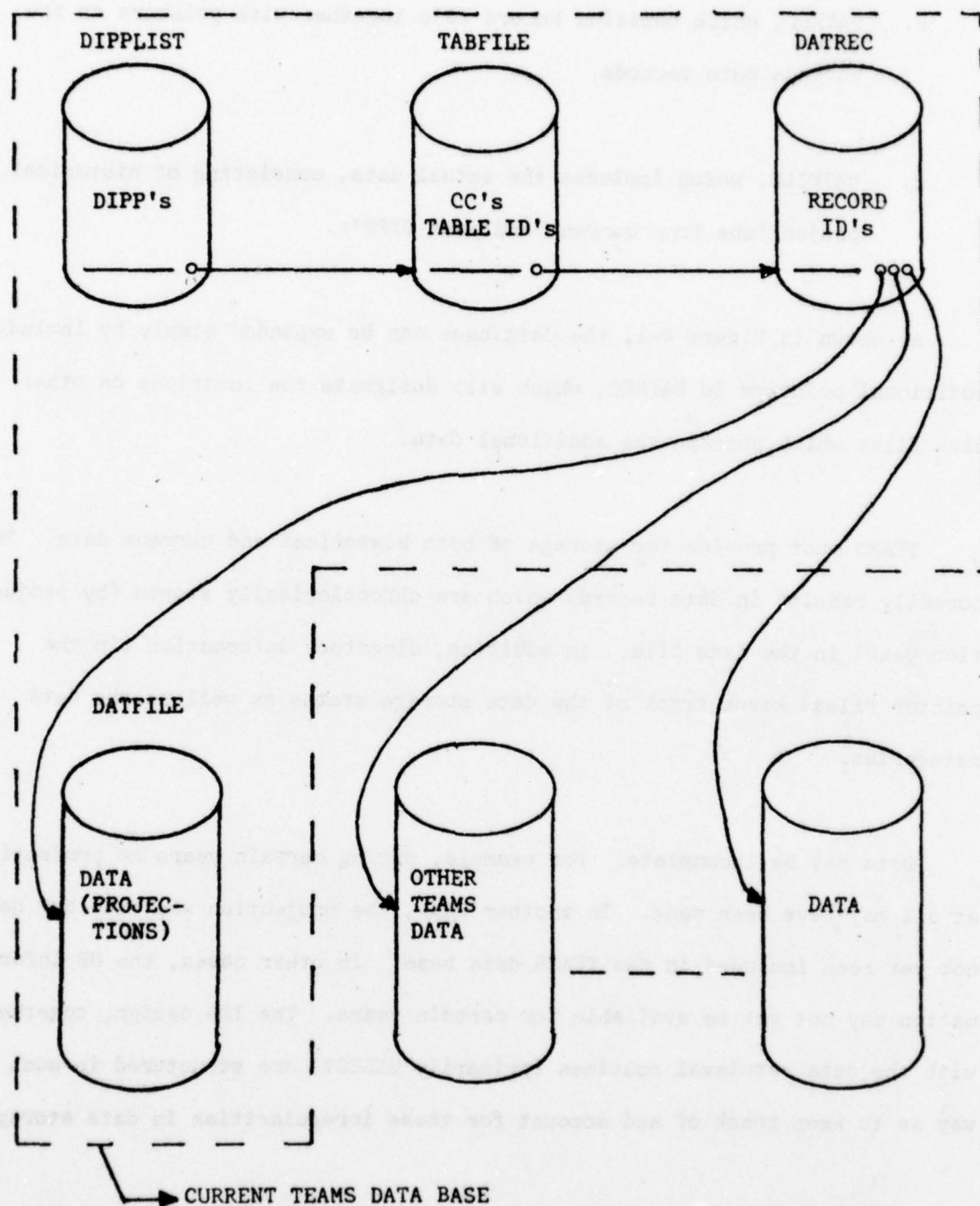
Figure 4-1 Expanding the TEAMS Data Base

4-4

storage when appropriate. A special Country Code allows projection data to be stored and retrieved by country. The Non-Soviet Warsaw Pact countries, for example, include more than one country in a single DIPP, and it is therefore desirable to provide a method for specifying the individual country to be retrieved.

Finally, it should be noted that the TDB has retained the basic DIPPOLS record fields, mnemonics, etc., whenever possible, so that there is a continuity of design between the two. The primary differences between the DIPPOLS and TEAMS data bases (aside from architecture) are the additions of Country Codes and a revised Table ID format.

4.2.   THE WORK AREA

Each TEAMS user is assigned a Work Area (WA), which functions as a notebook for storing information, making computations, and preparing reports. The WA remains unchanged from session to session, but can be erased at the command of the individual user. The WA is a file which is stored on disk, and which can be written, modified, and read by the user from a typewriter terminal; contents of the WA can also be directed to the line printer and the computer-driven plotter.

Initially, the WA is filled with zeros, except for a few "Housekeeping" records at the beginning. As space is used in the WA, the housekeeping (HK) records keep track of the amount of remaining space available; in addition,

they note the "garbage," or records which are no longer needed by the user. The WA contains four parts:

a.   The housekeeping records.

b.   SYMBOT, which is a list of pointers to the data stored in the WA, together with the names of the variables which the user has chosen.

c.   Freespace, which is that portion of the WA which is not currently in use.

d.   Data Storage, consisting of the information which the user has stored in the WA.

The user gives names to the various pieces of data which are to be stored in the WA. These pieces of data can take any of the following forms:

a.   Scalar. This is simply a single number, real or integer, positive or negative.

b.   Vector. A vector is a list of real or integer numbers.

c.   Text string. This is any information, written as textual material.

d.   Mergfile. This is a set of one or more projection records, including keys to identify the records.

e.   Statfile. A statfile is a mergfile which contains High, Low, and
      Best estimates, together with the most recent OB estimate for com-
      parison (if available). This format is used for statistical compu-
      tations.

f.   Plot format. This is a text string which contains parameters for
      the plotting routines. Since plot formats are somewhat lengthy,
      they may be stored in the WA and re-used when required, rather than
      requiring the user to type them in each time.

Integrity of the WA is maintained through frequent checks, of which the
following are examples:

a.   A system of redundant pointers helps to insure that the correct data
      are located. Specifically, pointers to the required data contain
      the location at which the data item is stored; in addition, the data
      item contains a reference to the original pointer, which is checked
      by the retrieval program.

b.   Instructions such as INHIBIT and LOCK serve to prevent inadvertent
      erasure of critical data. The CLEAR instruction requires confir-
      mation before it operates. In the extreme case, when data have been
      unintentionally lost, it will often be possible for the system
      programmer to restore the pointers to the data, since the data are
      not physically erased by the system; only the pointers are deleted.

4-7

c.   Before an assignment of data is made, TEAMS checks the WA to deter-
     mine whether the destination variable already exists, or whether the
     INHIBIT flag is on. This will help to prevent inadvertent destruc-
     tion of data.

d.   TEAMS always determines whether there is sufficient space in the WA
     for the requested operation. This serves to prevent the problems
     that would occur if an assignment were initiated but could not be
     completed.

To make the best use of the available space in the WA, a garbage col-
lection routine is called automatically whenever the remaining freespace is
inadequate for a requested operation. The purpose of a garbage collection
routine is to identify areas of storage which are not being used for data, and
which may be used to store new data. Since garbage collection is a relatively
time-consuming process, the routine is called only when needed. In addition,
since garbage collection is performed only when required, it will often be
possible for the system programmer to restore data which have been inadver-
tently erased.

The structure of the WA is such that no limits are imposed on the length
of variables, other than the limitation of the total size of the WA itself.
Thus, long texts, notes, and descriptive material, as well as numerical and
tabular information, can be stored in the WA. The length of the variable is
stored in the variable header.

Since the WA is stored on disk, it will not be destroyed by a power failure or by most other sources of system breakdown (other than physical damage to the disk).

The size of the WA is set by the system manager, and may be increased or decreased according to need.

4.3.   THE TEAMS COMMAND LANGUAGE

TEAMS includes a flexible command language to handle the acquisition of data from the data base, the computation of the TEAMS statistics and the creation of tables and plots that may be requested by the user. The syntax of the language is simple and easy to learn. The details and most of the technicalities of TEAMS are transparent to the user. Storage space in the Work Area is monitored by TEAMS; if there is not enough storage for the user's current operation, TEAMS informs him of that fact.

In addition to frequent integrity checks by the system, the user has two commands that prevent the inadvertent destruction of data. One is the INHIBIT command which prevents a change to or the erasure of any existing variable in the user's Work Area. The inhibit feature can be overridden by the user. The other command is the LOCK command which prevents a change or an erasure of a specific variable and cannot be overridden until the user enters an UNLOCK command.

Each time a computation or a selection from the TEAMS data base is made, the user must specify the name of a variable in which the result is to be stored. Variables can also be created by the direct user input of data using the ASSIGN command.

In order better to control WA storage, TEAMS has several commands that give the user information concerning the state of the WA. The command MAP causes the display of the basic WA storage parameters. The command VARIABLES lists the names of all variables currently in the WA.

For the casual user, TEAMS provides prompts to tell him what is expected next. For example, if a user types in the name of a command and then presses the carriage return, the system responds with a prompt that informs the user what is expected next. At any time, the user may type in "?" to request additional information. He may at any input type just a carriage return which will immediately return control to command input mode.

One of the key features of the TEAMS command language is the ease with which data can be put into tabular form. The user first creates a mergfile or statfile (using SELECT or STATFILE) and then uses the TABLE command, possibly requesting additional statistics. The table is then displayed at the terminal. He can also cause the same table to be printed on the line printer using the LTABLE command. The details for creating a batch job to produce the line printer output are handled by the system.

The EXECUTE statement allows the user to put together a command string that is stored in the WA and can be called for future use. The variables in the EXECUTE statement can be dummy variables and are set when the command string is called. This statement is most useful when a series of commands will be used often.

Another useful command is MODIFY. It enables the user to change parts of a vector, text string, mergfile or statfile to his specifications. Statistical computations can then be made on the altered data and the results compared with the original statistics. In this way, the analyst can study the effects of different predictions and arrive at a reasonable set of projections.

Another important feature of TEAMS is the ability to create plots with up to 10 curves on a plot. In the PLOT command, the user specifies the format of the plot (including paper size, length of axes, labels, etc.) and the variables to be plotted.

TEAMS provides statistical commands to assist in the evaluation of historical estimations. These are discussed in Section 6.

Table 4-1

TABLE OF TEAMS COMMANDS

Work Area Data Manipulation Commands

ASSIGN   Assigns variable names to data for input into TEAMS.

CLEAR   Clears the WA of all variables and data. Subject to
                        INHIBIT and LOCK.

COMBINE   Combines two variables together to produce a third.

ERASE   Erases a variable.

EXECUTE   Causes a command string that is stored in a text
                        string variable to be executed.

FETCH   The data portion of a mergfile logical record is

                        retrieved and stored in a new variable.

IF   Used for conditional execution of a command depending

                        on the existence of a variable in the WA.

IFNOT   The opposite of IF.

INHIBIT   Causes confirmation to be requested from the user

                        before any changes or erasures are made of any vari-

                        able in the WA.

LOCK   Prevents the contents of a variable from being

                        changed or destroyed.

MODIFY   Changes the contents of a variable according to user-

                        supplied criteria.

STOP   Terminates a TEAMS user session.

UNLOCK   Removes locking condition from a variable.

Inquiry Commands

LENGTH   Prints the length of a variable and number of WA

                        records used to store that variable.

MAP   Prints a summary of the major amounts of storage in

                        use in the WA and graphically displays the WA at the

                        terminal.

MENU   Prints a list of valid command names.

VARIABLES   Prints a list of the name and type of each variable

                        in the WA.

Display Commands

LPRINT   Causes the information in a variable to be printed on
                          the line printer.

LTABLE   Causes the contents of a mergfile or statfile, plus
                          any requested statistics, to be printed on the line
                          printer.

PLOT   Produces plots on the plotter as chosen by the user.

PRINT   Causes the information in a variable to be printed at
                          the terminal.

TABLE   Causes the contents of a mergfile or statfile plus
                          any requested statistics to be displayed at the
                          terminal.

Data Retrieval Commands

BESTOB   Finds the best (latest) known OB value or values for
                          a set of keys.

SELECT   Gathers logical records from the TEAMS data base or
                          from a mergfile according to a set of keys and places
                          them in a new mergfile.

STATFILE   Gathers logical records from the TEAMS data base

                            according to a set of keys and places them in a

                            statfile.

Statistical Commands

BIAS   Calculates the TEAMS bias statistics.

DRANK   Produces the descending ranks for the contents of a

                            vector variable.

DSORT   Produces a descending sort of a vector or mergfile.

ERROR   Calculates the TEAMS error statistic.

PDSTAT   Computes the mean and standard deviation for the data

                            corresponding to each prediction year in a mergfile.

PJSTAT   Computes the mean and standard deviation for data in

                            a vector or for the data in each logical record of a

                            mergfile.

SCORE   Computes the TEAMS scoring rule.

TAU   Computes the Kendall Tau statistic and the associated

                            probability.

UNCERTAINTY   Computes the TEAMS uncertainty statistic.

URANK   Produces the ascending ranks for the contents of a

                           vector variable.

USORT   Produces a descending sort of a vector or mergfile.

WERROR   Computes the TEAMS weighted error statistic.

SAMPLE TEAMS USER SESSION

This section illustrates several of the TEAMS commands and some typical output they produce.

RADC R&D TSS GCOS-SRH 8/14/77 AT 11.149 CHANNEL 3160

LOGON ID-QRSVBC34; 345810943067

PASSWORD--

SYSTEM ?FORT N

READY

*RUN TEAMS

      WELCOME TO TEAMS

COMMAND?   (This command simply lists all the available

  MENU   TEAMS commands and their necessary fields.)

--- COMMANDS ---

    ASSIGN   -DESTVAR-   -DATA OR SOURCEVAR-

    BESTOB   -DESTVAR-   -KEYS OR TSTG-

    BIAS   -DESTVAR-   -VECVAR OR STATF-   -PARM-

    CLEAR

    COMBINE   -DESTVAR-   -SOURCEVAR-   -SOURCEVAR-

    DRANK   -DESTVAR-   -VECVAR

    DSORT   -DESTVAR-   -VECVAR OR MERGF-

    ERASE   -VARNAME-

    EXECUTE   -AUXPARM-   -CMDTSTG-   -PARM...-

```
ERROR       -DESTVAR-        -VECVARWG OR STATF-      -PARM-

FETCH       -DESTVAR-        -PARM-      -MERGF-

IF          -DATA OR VARNAME-    -CMPCMD-

IFNOT       -DATA OR VARNAME-    -CMPCMD-

INHIBIT

LENGTH      -VARNAME-

LOCK        -VARNAME-

LPRINT      -VARNAME-

LTABLE      -MERGF OR STATF-

MAP

MENU

MODIFY      -DESTVAR-        -SOURCEVAR-     -RPCNSTG-

PDSTAT      -DESTVAR-        -MERGF-

PJSTAT      -DESTVAR-        -VECVAR OR MERGF-

PLOT        -PFVARNM-        -PTSTG-     -VARNAME...-

PRINT       -VARNAME-

SCORE       -DESTVAR-        -STATF OR VECVAR-    -VECVARWG-

SELECT      -DESTVAR-        -MERGF-     -KEYS OR TSTG-

STATFILE    -DESTVAR-        -KEYS OR TSTG-

STOP

TABLE       -MERGF OR STATF-      -STATNAME-

TAU         -DESTVAR-        -VECVAR OR MERGF-    -VECVAR OR MERGF-

UNCERTAINTY    -DESTVAR-         -VECVAR OR STATF-    -PARM-

UNLOCK      -VARNAME-

URANK       -DESTVAR-        -VECVAR-

USORT       -DESTVAR-        -VECVAR OR MERGF-
```

VARIABLES

WERROR   -DESTVAR-   -VECVAR OR STATF-

→ SELECT MF * '+U,SS,4B01H,GCJ,77'

'MF' NOW DEFINED

(A mergfile MF has been created with the high predictions made

in 1977 for weapon type GCJ.)

→ PRINT MF

U SS 4B01H GCJ 77 JULIETT 77/03/18

1973 15

22 24 26 28 30 34 36 38 36

34 36 38 36 30 28

→ STATFILE STF '+U,SS,4B01H,GCJ,71'

'STF' NOW DEFINED

(A statfile STF has been created. It contains low, best, high

and latest OB predictions and data.)

→ PRINT STF

U SS 4B01L GCJ 71 JULIETT 71/08/24

1967 15

15 15 17 20 19 21 23 24 27 28

30 35 32 30 26

U SS 4B01B GCJ 71 JULIETT 71/03/26

1967 15

15 15 17 20 20 23 25 26 30 30

34 37 33 32 30

U SS 4B01H GCJ 71 JULIETT 71/08/16

1967 15

15 15 17 20 22 25 28 30 33 33

35 40 35 35 33

U SS 4B01O GCJ 71 JULIETT *****

1967 11

16 17 20 21 22 23 26 28 30 31

33

(The user now wishes to obtain a printout of STF in tabular and geographical form. He realizes that he must get similar output for statfiles he will create in the future. He first uses the ASSIGN command to create a command string CMD. PLFOR is an already existing plot format variable.)

→ ASSIGN CMD   "LTAB %1 'UNCERT' PLOT PLFOR 'V' %1"

'CMD' NOW DEFINED

(He then uses the EXECUTE command to create the table and plot for STF.)

→ EXECUTE 1 CMD STF

DOING LTAB

LINE PRINTER TABLE OUTPUT SUBMITTED

DOING PLOT

PLOTTER OUTPUT SUBMITTED

(The variable STF replaces the %1 in CMD when CMD is executing.
The table produced will be similar to Table 5-1.)

The plot produced by the TEAMS plotting routine is similar to
Figure 5-1.

The user now lists the variable names present in his Work Area
and obtains information on the storage use in his Work Area.

→ VARIABLES

TYPE	NAME
TEXT STRING	TEMP
TEXT STRING	HI
MERGFILE	GORMPH
VECTOR	VECTOR2
MERGFILE	MORFIC
MERGFILE	RAND
VECTOR	BIASVECT
MERGFILE	THING
PLOT FORMAT	PLFOR
TEXT STRING	CMD
MERGFILE	MF
STATFILE	STF

U SS 4B01 GCD 1971

	LOW	BEST	HIGH	BESTOB	UNCERT
1967	15	15	15	16	.00
1968	15	15	15	16	.00
1969	17	17	17	18	.00
1970	20	20	20	21	.00
1971	19	20	22	22	.14
1972	21	23	25	23	.16
1973	23	25	28	26	.18
1974	24	26	30	28	.20
1975	27	30	33	30	.18
1976	28	30	33	31	.15
1977	30	34	35	33	.14
1978	35	37	40		.13
1979	32	33	35		.04
1980	30	32	35		.14
1981	26	30	33		.21

Table 5-1   Sample Line-Printer Table

Figure 5-1 Sample Plotter Output

(Number of Planes)

40 38 36 34 32 30 28 26 24 22 20 18 16 14

67 68 69 70 71 72 73 74 75 76 77 78 79 80 81

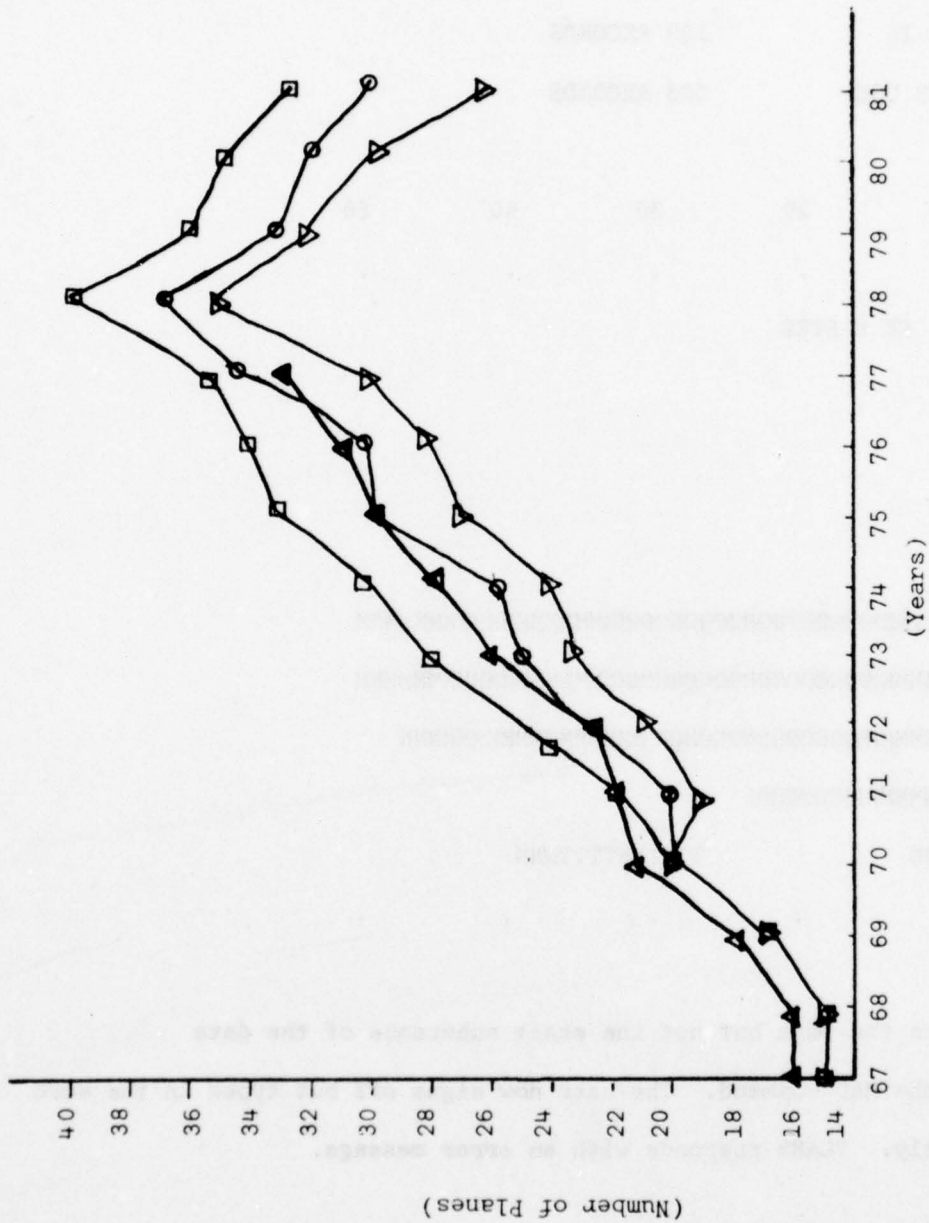(Years)

→ MAP

```
        DIRECTORY CONTAINS        25 ENTRIES

        FREESPACE IS             139 RECORDS

        DATA&NAMES USE           333 RECORDS


    1         10        20        30        40        50

    .         .         .         .         .         .

  1HHH EE    E  EE E EEEE

 51

101

151

201

251          VVNNMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM

301MMMMMMMMMMMMMMMMMMNNVVNNMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM

351MMMMMMMMMMMMMMMMMMMMMMMMMMMMNN MMMMMMMMMMMMMMNNN

401 VVVNNMMMMMMMMMMMMMMMMMNN

451   TTTTTNNN            TTTTNNTTTTNNN

    .         .         .         .         .         .
```

This represents the form but not the exact substance of the data
output under the MAP command. The user now signs off but types in the word
STOP incorrectly. TEAMS responds with an error message.

→ STPO

    INCORRECT COMMAND NAME

This time he corrects his mistake.

```
STOP

SIGN-OFF COMPLETE

*BYE

**cost:  $ 5.41 to date:  $ 122.26 = 2%

**on at 20.149 - off at 21.359 on 8/14/77
```

SECTION 6

STATISTICAL METHODS

In this section, statistical methods used in TEAMS programs are briefly described. TEAMS statistics are designed to provide a quantitative tool for the analysis of past predictions. More specifically, TEAMS statistics may be utilized in the following ways:

a. The detection and quantification of past error are important factors in quality control. Several of the statistics are measures of error that can be compared across different weapon systems for identifying trouble spots. They are designed to locate the areas in which the more serious errors occurred, to provide the basis for a review of methodology.

b. Measurements of the accuracy of predictions are provided as feedback which may be used for self-rating, and to detect areas in which a methodological review may be required.

c. A measurement of uncertainty is provided as a means for determining the degree of confidence that may be attributed to a projection.

d. Trends in data may be identified for analysis.

e.   Combined with access to predictions made in the past and the most
     recent OB values, the statistics should lead to a better under-
     standing of the data and to improved predictions.

The following TEAMS statistics are described in this section:

a.   The Scoring Rule (6.1.).

b.   Mean and Standard Deviation (6.2.).

c.   Uncertainty, Bias, Error and Weighted Error (6.3.).

d.   Ranking and Sorting (6.4.).

e.   Kendall Tau Statistic (6.5.).

6.1.   THE SCORING RULE

In order to provide analysts and managers with a fair and flexible rule
by which scores could be assigned to projection predictions, a scoring rule
was designed to meet the following criteria:

(a) Scores should take on values between 0 and 1.

(b) The rule should be mathematically elementary so that the analyst
     could understand it, have confidence in it, and be able to compute
     scores using a hand calculator.

(c) The rule should be "proper" in the sense that it would encourage the
     analyst to report what he believes.

(d) The rule should utilize only the low, best, high and OB data and the
confidence that the analyst ascribes to the interval between his low
and high estimates. These are the statistics that are well under-
stood by the analyst and that are maintained in the historical
files.

(e) Scores for predictions of different weapon systems should be com-
parable.

(f) The scores actually obtained from the rule should yield a reasonable
set of numeric values. A very precise projection which is confirmed
should be given a score close to 1, while an imprecise projection,
or one that is badly disconfirmed, should be given a score near
zero.

(g) Analysts often give different meanings to the projection statistics
that they report. For example, one analyst might believe that his
best estimate represents his true belief concerning the event most
likely to occur, while another analyst can only be sure of his high
and low estimates and reports a best only because he is required to
do so. In order to accommodate such different points of view, a
certain flexibility must be built into the scoring rule. This we
have accomplished by allowing the analyst to specify weights of
importance for various terms in the rule.

(h) A scoring rule is designed to indicate the presence of error and the magnitude of the error. It cannot point to the source of the error. This is the task of the analyst.

(i) The scoring rule is not intended to be a rating system for individual analysts. There is a wide diversity in the difficulty of the various tasks performed, with the result that the more difficult projections will almost inevitably receive lower scores. If these scores are then used for rating individual analysts, the result will be to discourage them from undertaking the more difficult tasks, in order to avoid the low ratings that they are likely to receive. It is best to regard low scores as evidence that the projections were difficult, the data were faulty, or some other source of error was present.

In the following sections, L, B, H and ω (omega) stand for low, best, high and OB values respectively. C is the chance that the real value will fall in the L to H interval. C is called the confidence coefficient and ranges from 0 to 1.

Many scoring rules were considered in the course of this investigation. The scoring rule described in Sections 6.1.1. and 6.1.2. is the one that PAR feels is best suited to the purpose of the DIA analysts and provides sufficient flexibility for use in different situations. Section 6.1.3. describes labeling and sorting errors and their use. In Section 6.1.4. we discuss the

logarithmic scoring rule that was considered during the first phase of this project; it is included as an alternative approach but has not been described in TEAMS system specifications.

6.1.1.   The Rule and Examples

Define the factors

$$
f_1 = \begin{cases} \dfrac{B - \omega}{2B+1-\omega} & \text{if } L \le \omega \le B \\[4mm] \dfrac{\omega-B}{\omega+1} & \text{if } B < \omega \le H \end{cases}
$$

$$
f_2 = \frac{H-L}{H+1}
$$

$$
f_3 = \begin{cases} \dfrac{B-L}{H-L+1} & \text{if } L \le \omega \le B \\[4mm] \dfrac{H-B}{H-L+1} & \text{if } B < \omega \le H \end{cases}
$$

$$
f_4 = 1 - C.
$$

f_1 is a measure of the accuracy of the point estimate ω. f_2 is essentially the TEAMS uncertainty statistic. f_1 and f_3 are not defined for $\omega < L$ or $\omega > H$. f_3 is a measure of relative uncertainty, i.e. how does B-L compare with H-B when $L \le \omega \le B$? The "1" is placed in the denominators of f_1, f_2 and f_3 to avoid the possibility of division by zero.

The score $S(\omega)$, for $L \leq \omega \leq H$ is then defined as

$$S(\omega) = 1 - \left[\frac{af_1 + bf_2^2 + cf_3^2 + df_4^2}{a + b + c + d}\right],$$

where a, b, c, d are the weights or measures of importance attached to the factors f_1, f_2, f_3, f_4 respectively. The squares of f_2, f_3 and f_4 are more often similar in size to f_1 and so are used in the definition. Since the numbers f_1, f_2, f_3 and f_4 all lie between 0 and 1, $S(\omega)$ is a number between 0 and 1.

To define the score in the regions outside $L \leq \omega \leq H$ let

$\alpha = S(L)$ (= the score when $\omega = L$),

$\beta = S(H)$ (= the score when $\omega = H$)

and $\qquad X = \dfrac{b(f_2) + d(f_4)}{b + d}$

The score $S(\omega)$ is then defined for $\omega < L$ or $\omega > H$ by

$$S(\omega) = \begin{cases} \dfrac{\alpha}{1+\left(\frac{L-\omega}{H-L}\right)\frac{1}{X}} & \text{if } \omega < L \\[4ex] \dfrac{\beta}{1+\left(\frac{\omega-H}{H-L}\right)\frac{1}{X}} & \text{if } \omega > H \end{cases}$$

It is assumed above that $L \leq H$; when this condition does not hold, the two values should be interchanged.

Examples:

Suppose that an analyst makes a prediction of L=6, B=8, H=10 and an OB score of ω=7 is found. Suppose that he has chosen the weights as a=4, b=4, c=1 and d=1 and his confidence was 75%, i.e. C=.75. Then ω is in the low, high range, and he computes:

$$f_1 = \frac{8 - 7}{16 + 1 - 7} = .1$$

$$f_2 = \frac{10 - 6}{10 + 1} = .364$$

$$f_3 = \frac{8 - 6}{10 - 6 + 1} = .4$$

$$f_4 = 1 - .75 = .25$$

$$S(7) = 1 - [\frac{4(.1) + 4(.364)^2 + 1(.4)^2 + 1(.25)^2}{4 + 4 + 1 + 1}]$$

$$= .884.$$

Suppose now that an OB figure of ω = 12 was obtained. First the analyst must compute β. To do this note that f_2, f_3, f_4 are as above and

$$f_1 = \frac{10 - 8}{10 + 1} = .182.$$

Then,

$$\beta = S(10) = 1 - [\frac{4(.182) + 4(.364)^2 + 1(.4)^2 + 1(.25)^2}{4 + 4 + 1 + 1}]$$

$$= .852.$$

Next, X must be computed:

$$X = \frac{4(.364) + 1(.25)}{4 + 1} = 0.341.$$

Finally,

$$S(12) = \frac{.852}{1+(\frac{12 - 10}{10 - 6})\frac{1}{.341}} = 0.345$$

6.1.2.   Additional Properties of the Scoring Rule

(a) The scoring system is quite flexible. For instance, by changing

the weights (a, b, c and d) the analyst or manager can "tune-up"

the scoring system so that the output of scores more closely

matches a set of desired ones, or to adjust for the importance of

each of the measures. For example, if the accuracy of the best

estimate is of more importance than the degree of uncertainty, a

possible choice of weights would be a a = 6, b = 4, c = d = 1.

Once weights are decided upon for a particular set of weapon

systems they should be fixed.

(b) Analysts are encouraged to report what they believe. In the range

$L \leq \omega \leq H$, scores decrease as $B - \omega$ increases, as $H - L$ increases.

or as C decreases. Outside of the range $L \leq \omega \leq H$ scores increase as $H - L$ increases or as C increases.

(c) Scores are comparable from weapon system to weapon system. If the weights are fixed, then the score for a set L, B, H, ω of data is approximately equal to the score for the set kL, kB, kH, kω where k is any positive constant. (In fact, if the ones were removed from the denominators of f_1, f_2 and f_3 then the two scores would be exactly equal.)

(d) The scoring rule has properties of symmetry. If $H - B = B - L$, then the score at $B - h$ equals the score at $B + h$ for any h such that $B - h \geq 0$, i.e.

$$S(B - h) = S(B + h) \text{ if } B - h \geq 0.$$

If the weight c is chosen to be zero then $S(B - h) = S(B + h)$ as long as $B - h$ and $B + h$ are both contained in the low, high range.

(e) If in some data set $L > H$, the values of L and H are interchanged and the score is computed as before.

(f) The scoring rule is accessed by a TEAMS command SCORE. For further details, see the User's Guide to TEAMS.

Sample data for the scoring rule are contained in Table 6-1.

6.1.3.   Sorting and Labeling Error

In order to discuss sorting and labeling error, a simple error score $E(\omega)$ is introduced that depends only on whether ω is in the low, high range and on the confidence coefficient C. More specifically, given a prediction L, B, H and an OB value ω, we define

$$E(\omega) = \begin{cases} 100\,(1 - C)^2 & \text{if } L \leq \omega \leq H \\[2em] 100\,C^2 & \text{if } L < \omega \text{ or } \omega > H. \end{cases}$$

For example, if L = 10, B = 15, H = 20 and ω = 17 and C = .80, then the error is

$$E(17) = 100(1 - .80)^2 = 4.$$

If ω = 21 (or 22 or 307), on the other hand, the error score is

$$E(21) = 100(.80)^2 = 64.$$

Consider now a collection of n predictions all using the same confidence coefficient C and suppose that x of them are "hits," that is, ω lies in the low-to-high range x times. The <u>average error</u> for this group of predictions is then defined as

6-10

LOW	BEST	HIGH	BOB	SCORE
10.	20.	35.	5.	0.4939
10.	20.	35.	8.	0.5804
10.	20.	35.	10.	0.6570
10.	20.	35.	12.	0.6757
10.	20.	35.	15.	0.7091
10.	20.	35.	17.	0.7361
10.	20.	35.	18.	0.7513
10.	20.	35.	19.	0.7679
10.	20.	35.	20.	0.7861
10.	20.	35.	21.	0.7494
10.	20.	35.	23.	0.7176
10.	20.	35.	25.	0.6906
10.	20.	35.	28.	0.6572
10.	20.	35.	30.	0.6385
10.	20.	35.	32.	0.6221
10.	20.	35.	35.	0.6009
10.	20.	35.	37.	0.5308
10.	20.	35.	40.	0.4517
10.	20.	35.	47.	0.3352
10.	20.	35.	55.	0.2589
100.	130.	200.	90.	0.6633
100.	140.	200.	90.	0.6464
100.	150.	200.	90.	0.6301
100.	130.	200.	120.	0.8576
100.	140.	200.	120.	0.8294
100.	150.	200.	120.	0.8039
100.	160.	200.	120.	0.7798
100.	130.	220.	120.	0.8413
100.	130.	250.	120.	0.8186
100.	130.	300.	120.	0.7866
100.	130.	200.	125.	0.8712
100.	130.	200.	129.	0.8829
100.	130.	200.	133.	0.8378
100.	130.	200.	131.	0.8437
100.	130.	200.	139.	0.8210
100.	130.	200.	147.	0.8008
100.	130.	200.	156.	0.7805
100.	130.	200.	165.	0.7624
100.	130.	200.	178.	0.7394
100.	130.	200.	187.	0.7254
100.	130.	200.	199.	0.7087
100.	130.	200.	201.	0.6920
100.	130.	200.	220.	0.4891
100.	130.	200.	300.	0.2189
700.	800.	900.	789.	0.9439
700.	800.	900.	793.	0.9458
700.	800.	900.	900.	0.9049

Table 6-1 Sample Data for the Scoring Rule with C = .75, a = 4, b = 4, c = 1, and d = 1

$$E_A = \frac{x[100(1 - C)^2] + [(n - x)(100 \, C^2)]}{n} \, .$$

The __sorting error__ is defined as

$$E_S = 100 \, \frac{x}{n} \, (1 - \frac{x}{n})$$

The labeling error is defined as

$$E_L = 100(\frac{x}{n} - C)^2 .$$

These errors are connected by the equation

$$E_A = E_S + E_L \, ,$$

that is, the average error equals the sum of the sorting error and the labeling error.

Proof:

$$
\begin{aligned}
E_S + E_L &= 100 \, \frac{x}{n}(1 - \frac{x}{n}) + 100(\frac{x}{n} - C)^2 \\[2mm]
&= 100 \, [\frac{x}{n} - (\frac{x}{n})^2 + (\frac{x}{n})^2 - 2C\frac{x}{n} + C^2] \\[2mm]
&= 100 \, [(1 - 2C)\frac{x}{n} + C^2]
\end{aligned}
$$

$$E_A \quad = \quad 100 \left[\frac{x}{n} - \frac{2x}{n}C + \frac{x}{n}C^2 + (1 - \frac{x}{n})C^2\right]$$

$$= \quad 100 \left[\frac{x}{n} - \frac{2x}{n}C + \frac{x}{n}C^2 + C^2 - \frac{x}{n}C^2\right]$$

$$= \quad 100 \left[(1 - 2C)\frac{x}{n} + C^2\right] .$$

The sorting error E_S can be thought of as a measure of the degree of certainty the analyst has concerning a specific weapon system or group of systems. If knowledge concerning the system is comprehensive and detailed, then the analyst would sort predictions into groups that are almost all true (so that $1 - \frac{x}{n}$ will be close to zero). In this case, the sorting error will be low. If knowledge is incomplete, then the sorting error will be large.

If an analyst correctly values his knowledge, the fraction of his forecasts which fall within the low-to-high range specified should be close to C. Labeling error is a mathematical measure of how correctly an analyst values his knowledge. Large labeling error indicates that the analyst's confidence coefficient might be in error or that the low, high interval spreads are incorrect. Low labeling error indicates that knowledge is being correctly reported.

Example. Suppose that for a group of 25 aircraft predictions with C = .75, ω lies in the low, high range 20 times and outside the range 5 times.

Thus, $x = 20$, $n = 25$ and

$$E_A = \frac{20[100(.25)^2] + 5[100(.75)^2]}{25} = 16.25$$

$$E_S = 100(\frac{20}{25})(\frac{5}{25}) = 16$$

$$E_L = 100(\frac{20}{25} - .75)^2 = .25$$

In this example, the labeling error is quite low and hence C is a very good confidence coefficient and the choices of L and H were, on the average, appropriate.

6.1.4.   Proper Scoring Rules

A mathematically proper scoring rule is one in which the expected value of the score cannot be increased by "fudging"; that is, the score is maximized when the analyst reports exactly what he believes. In this situation, the analyst must report a complete set of probabilities for a feasible range of values. In the DIA-DE-1 environment, only low, best, high figures and a confidence coefficient are given, so that a probability distribution $P(\omega)$ must be fitted to the data in order to construct a proper scoring rule. In the course of our research, this was accomplished as follows.

If $2L - B \geq 0$:

$$P(\omega) = \frac{1}{2(H-L)(B-L)}(B - 2L + \omega) \qquad \text{for } 2L - B \leq \omega \leq B \text{ or } H = B$$

$$= \frac{1}{2(H-L)(H-B)}(2H - B - \omega) \qquad \text{for } B \leq \omega \leq 2H - B \text{ or } L = B$$

$$= 0 \qquad \text{otherwise.}$$

If $2L - B < 0$:

$$p(\omega) = \frac{\dfrac{(B - 2L + \omega)}{2(B - L)}}{H - \dfrac{B}{4}\left(\dfrac{B}{B - L}\right)} \qquad \text{for } 0 \leq \omega \leq B \text{ or } H = B$$

$$= \frac{\dfrac{(2H - B - \omega)}{2(H - B)}}{H - \dfrac{B}{4}\left(\dfrac{B}{B - L}\right)} \qquad \text{for } B \leq \omega \leq 2H - B$$

$$= 0 \qquad \text{otherwise.}$$

A confidence coefficient of $C = .75$ was used and $p(\omega)$ was constructed so that the probability that $L \leq \omega \leq H$ equals .75 in the case that $2L - B \geq 0$. See Figure 6-1 and 6-2 for graphs of $p(\omega)$.

The logarithmic scoring rule, which is known to be proper, was most suitable in this context. The score is then given by the expression

$$\text{Score} = \ln(p(\omega)).$$

This rule has the disadvantage that if $p(\omega) = 0$, Score $= -\infty$ and so instead a truncated logarithmic function was applied to $p(\omega)$ to obtain the score:
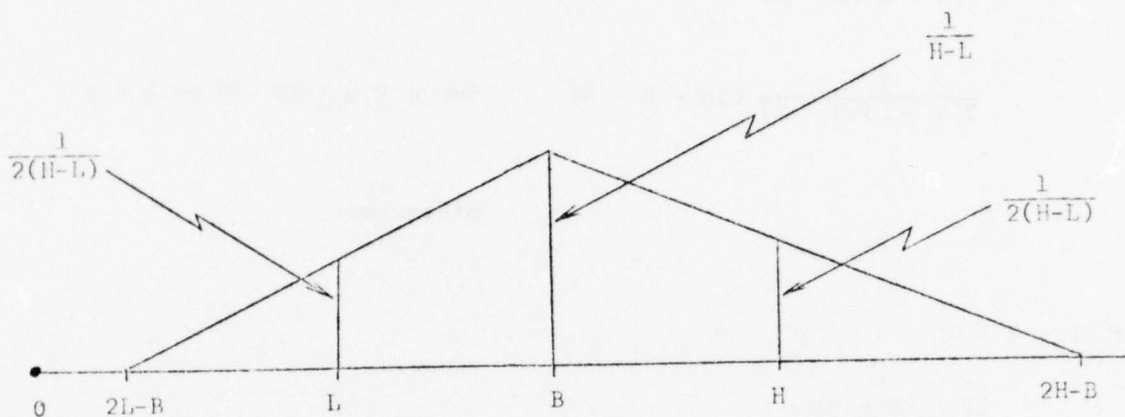
6-15

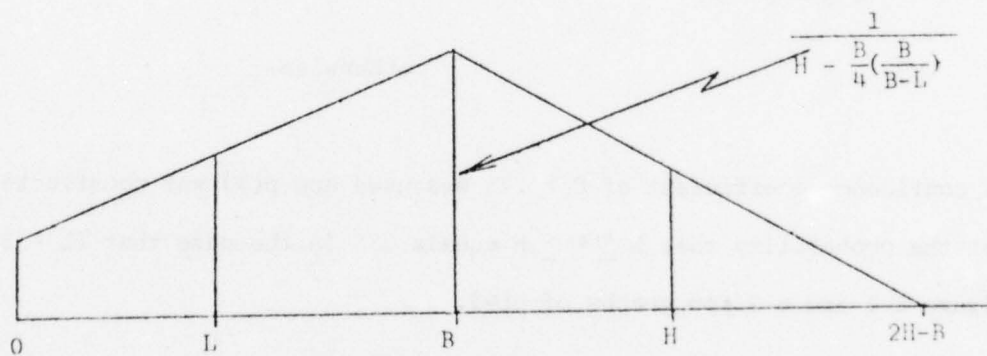Figure 6-1   p(ω) for the Case Where 2L - B $\geq$ 0

Figure 6-2   p(ω) for the Case Where 2L - B < 0

$$\text{Score}(\omega) = \ln(\max(M, P(\omega))) - \ln(M).$$

The constant M may be chosen to be a very small positive number; M = .001 and M = .0001 were used in our tests. This rule does not allow for comparisons across different weapon systems. This can be accomplished, however, by dividing the score by the upper bound U of possible scores:

$$U = \ln\left(\frac{1}{H - L}\right) - \ln M \qquad \text{for } 2L - B \geq 0$$

$$= \ln \frac{1}{H - \frac{B}{4}\left(\frac{B}{B - L}\right)} - \ln M \qquad \text{for } 2L - B < 0.$$

The normalized score, suitable for comparisons, is then given by

$$N(\omega) = \frac{\text{score}(\omega)}{U}$$

$$= \frac{\ln(\max(p(\omega), M) - \ln(M)}{U}.$$

Several disadvantages of this scoring system were noted as follows:

(a) Scores are rather high.

(b) An arbitrary probability distribution has been assumed that "looks good." It is not clear that in fact it is appropriate.

6-17

(c) The constant M in the term max $(M, p^{(\omega)})$ makes the rule not formally proper. However, this term is necessary to avoid scores of $-\infty$.

(d) If $H - L > 1000$, then the constant $M = .001$ gives 0 for all scores. One can correct this by using $M = .0001$, say, but then scores tend to be even higher.

For these reasons, it was decided not to recommend the use of this rule in TEAMS. A set of sample data using this rule can be found in Table 6-2.

6.2.   MEAN AND STANDARD DEVIATION

Mean and standard deviation are two useful statistics for the analysis of sets of data. If $\{x_1, x_2, \ldots, x_n\}$ is a set of n data elements, the mean or average $\overline{x}$ is defined by

$$\overline{x} = \sum_{i=1}^{n} x_i \, ,$$

and the standard deviation is defined by

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2}$$

The TEAMS commands PDSTAT and PJSTAT perform these computations for vectors and for mergfiles or statfiles. In the latter case the user has the option

				M = .001		M = .0001	
L	B	H	OB	Log Score	Normalized	Log Score	Normalized
10.	20.	35.	5.	2.3026	0.6242	4.6052	0.7686
10.	20.	35.	8.	2.7726	0.7516	5.0752	0.8471
10.	20.	35.	10.	2.9957	0.8121	5.2983	0.8843
10.	20.	35.	12.	3.1781	0.8615	5.4806	0.9147
10.	20.	35.	15.	3.4012	0.9220	5.7038	0.9520
10.	20.	35.	17.	3.5264	0.9559	5.8289	0.9729
10.	20.	35.	18.	3.5835	0.9714	5.8861	0.9824
10.	20.	35.	19.	3.6376	0.9861	5.9402	0.9914
10.	20.	35.	20.	3.6889	1.0000	5.9915	1.0000
10.	20.	35.	21.	3.6550	0.9908	5.9576	0.9943
10.	20.	35.	23.	3.5835	0.9714	5.8861	0.9824
10.	20.	35.	26.	3.4657	0.9395	5.7683	0.9628
10.	20.	35.	29.	3.3322	0.9033	5.6348	0.9405
10.	20.	35.	32.	3.1781	0.8615	5.4806	0.9147
10.	20.	35.	35.	2.9957	0.8121	5.2983	0.8843
10.	20.	35.	37.	2.8526	0.7733	5.1552	0.8604
10.	20.	35.	40.	2.5903	0.7022	4.8929	0.8166
10.	20.	35.	47.	1.3863	0.3758	3.6889	0.6157
10.	20.	35.	55.	0.	0.	0.	0.
100.	130.	200.	90.	1.2040	0.5229	3.5066	0.7614
100.	140.	200.	90.	1.3218	0.5740	3.6243	0.7870
100.	150.	200.	90.	1.3863	0.6021	3.6889	0.8010
100.	130.	200.	120.	2.1203	0.9208	4.4228	0.9604
100.	140.	200.	120.	2.0149	0.8751	4.3175	0.9375
100.	150.	200.	120.	1.9459	0.8451	4.2485	0.9225
100.	160.	200.	120.	1.8971	0.8239	4.1997	0.9120
100.	130.	220.	120.	1.9379	0.9140	4.2405	0.9588
100.	130.	250.	120.	1.7148	0.9039	4.0174	0.9566
100.	130.	300.	120.	1.4271	0.8867	3.7297	0.9534
100.	130.	200.	125.	2.2156	0.9622	4.5182	0.9811
100.	130.	200.	129.	2.2858	0.9927	4.5884	0.9964
100.	130.	200.	133.	2.2809	0.9906	4.5835	0.9953
100.	130.	200.	131.	2.2954	0.9969	4.5980	0.9984
100.	130.	200.	139.	2.2361	0.9711	4.5387	0.9856
100.	130.	200.	147.	2.1731	0.9438	4.4757	0.9719
100.	130.	200.	156.	2.0971	0.9108	4.3997	0.9554
100.	130.	200.	165.	2.0149	0.8751	4.3175	0.9375
100.	130.	200.	178.	1.8827	0.8177	4.1853	0.9088
100.	130.	200.	187.	1.7798	0.7730	4.0824	0.8865
100.	130.	200.	199.	1.6236	0.7051	3.9262	0.8526
100.	130.	200.	201.	1.5950	0.6927	3.8976	0.8464
100.	130.	200.	220.	1.2730	0.5528	3.5756	0.7764
100.	130.	200.	300.	0.	0.	0.	0.
700.	800.	900.	789.	1.5529	0.9649	3.8555	0.9855
700.	800.	900.	793.	1.5738	0.9779	3.8764	0.9909
700.	800.	900.	900.	0.9163	0.5693	3.2189	0.8228

Table 6-2 The Logarithmic Scoring Rule

of computing the mean and standard deviation for each logical record of the mergfile or statfile or for the data associated with each prediction year in the mergfile or statfile.

6.3.   UNCERTAINTY, BIAS, ERROR AND WEIGHTED ERROR

To define these statistics we let L, B, H and ω stand for low, best, high and order of battle estimates respectively. Each statistic is then defined in two forms: individual and group.

If L and H are the low and high predictions made for some prediction year then the (individual) uncertainty U is defined by

$$U = \frac{|H - L|}{\max(H,L)}$$

The absolute value in the numerator and the maximum in the denominator allow for the possibility that L > H. The numerator serves to measure the uncertainty and the denominator is a weighting factor which makes it possible to compare for different years or weapon systems.

If G is a group of L and H estimates then the group uncertainty U_G is defined as

$$U_G = \frac{\sum_G |H - L|}{\sum_G \max(H,L)}$$

Again, U_G is weighted so as to be comparable across different groups and U_G can be used to test for groups of weapons for which uncertainty was greatest. It is then up to the estimator to determine the cause.

If G is a group of L and H predictions made in the same projection, then U_G is the ratio of the area between the curves in Figure 6-3 to the area below the upper curve.

For a set of L, B, H and ω data, the TEAMS bias statistic BI is defined by

$$BI = \frac{B - \omega}{|H - L|}$$

For groups G, bias is defined by

$$BI_G = \frac{\sum\limits_{G} B - \omega}{\sum\limits_{G} |H - L|}$$

Again, the denominator is a weighting factor, making it possible to compare different biases. Unlike uncertainty, the terms in the numerator may be either positive or negative. Thus over a long enough period BI_G could tend to zero.

QUANTITY

H
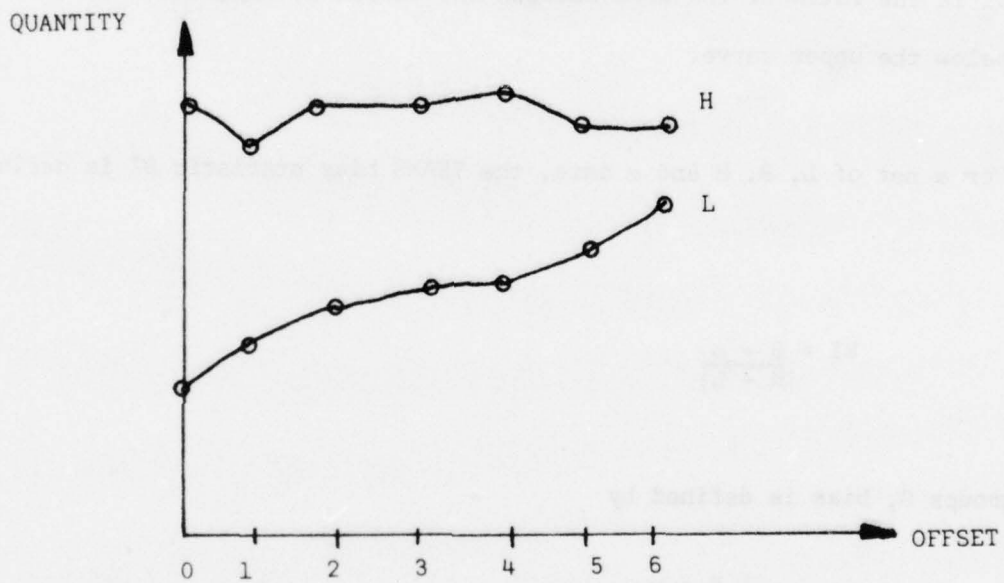
L

OFFSET

0   1   2   3   4   5   6

Figure 6-3

Example of High and Low Estimates Over Seven Years

The error E and weighted error WE for individual sets of B and ω data are defined by

$$E = |B - \omega|$$

and

$$WE = \frac{|B - \omega|}{\omega} .$$

For a group G of B, ω data these statistics are defined by

$$E_G = \sum_G |B - \omega|$$

$$WE_G = \frac{\sum_G |B - \omega|}{\sum_G \omega}$$

Error is simply the absolute difference between best estimates and ω values while weighted error is error normalized by ω so that it can be compared across different weapon systems.

6.4.   RANKING AND SORTING

TEAMS provides the capability for ranking and sorting lists of data through the following commands:

a.   DRANK, a high-to-low ranking routine.

b.   URANK, a low-to-high ranking routine.

6-23

c.   DSORT, a high-to-low sorting routine.

    d.   USORT, a low-to-high sorting routine.

A statistic, such as Bias or Error, can be computed for a set of groups, and
those groups can then be ranked or sorted according to the values of the
corresponding group statistics. In this way, groups or areas with highest
uncertainty, bias, error, etc., can be identified.

6.5.   KENDALL TAU STATISTIC

    The Kendall Tau correlation is a measure of the degree of association
between two variables. The use of the TEAMS TAU command returns the value
of the Tau statistic and a probability that has the following interpretation:
Under the hypothesis that there is no association between the variables, a
value of the Tau statistic greater than or equal to Tau is obtained (less
than or equal to Tau if Tau is negative).

    The Tau statistic can be used to answer questions such as the follow-
ing: Does error decrease with a decrease in uncertainty? Has error for a
particular weapon system decreased over time? Tau can also be used to check
for trends in data over time.

PROJECTION METHODS

To provide information required for design of the TEAMS user interface, interviews were conducted with DE-1 estimators and managers for the purpose of learning the terminology, methods, and general background of the personnel who will be using TEAMS. This section, which describes methods currently in use by DE-1, is based upon the interviews. In addition, some more general remarks are included concerning the sources of uncertainty which enter into long-range intelligence estimates.

The first subsection will provide a review of methods used in developing and presenting DE-1 estimates. The second subsection briefly describes some statistical methods which have been proposed for use in estimative intelligence. This is followed by suggestions derived from interviews with DE-1 personnel for the development of files containing annotations on projection methodology. Next, problem areas in estimative intelligence are reviewed, with particular application to sources of errors in DE-1 estimates. A further discussion of the communication of uncertainty is included.

7.1.   REVIEW OF PROJECTION METHODS

Estimative methods used by DE-1 have been summarized as follows: "The technique used in preparing the DIPP is to use the best available documentary intelligence and the best available expertise in arriving at a firm estimative

position and then, as appropriate, to describe or bound the limits of uncertainty inherent in these projections." Thus a Best estimate is developed; following this, a High and a Low estimate are formed, to indicate the degree of uncertainty of the projection. In practice, some estimators first select a High and a Low estimate, and then choose a Best estimate at a point between these extremes.

The range between High and Low figures is selected in such a way as to reflect the estimator's expectation that the actual figure will lie between these extremes in three out of four cases. Thus the High and Low figures represent a range in which the probability is 0.75 that the actual figure will lie between them, if the range has been estimated correctly. This probability represents a judgment by the estimator, not the result of a mathematical or statistical process; for this reason, it cannot be regarded as a mathematically precise figure.

No particular probability is ascribed to the Best estimate, since the degree of certainty will vary from one weapon system to the next, and will depend on the quantity and quality of information available, distance in time, and other factors. Although the expression "Best" suggests that this value is more probable than any other, some estimators state that all values in the range from Low to High are equally likely. In any case, there is no well-defined probability distribution of the values from Low to High.

There is some disagreement concerning the meaning of "Low" and "High." In certain instances, the Low value is greater than the High value; in such

instances it should be understood that a low level of technology has been projected, in which a semiobsolete system has been retained in greater numbers than would be retained under a high level of technology. As a matter of policy, such cases are now being regarded as anomalies, and in future projections, the Low figure will always be less than or equal to the High figure.

It might be better, however, to consider the Low table and the High table each to represent a consistent model of future conditions in which the Low estimates show the lowest reasonable level of effort that can be expected; similarly, the High estimates would show the highest levels. In such an approach, the actual numbers of small, obsolete, or low-technology systems would be greater in the Low tables than in the High tables. Such an approach seems to have been used in the past, producing the "anomalies" noted.

Methods used in producing DE-1 projections do not appear to be consistent among the estimators. One example of inconsistency was noted in the manner in which projections were developed: some persons prefer to work on Best estimates first, while others develop the Low and High estimates first. In some cases, estimates for individual weapon systems are formed first, while in other cases, estimates for classes of weapon systems are developed first. Although some general guidance and training are provided, estimators appear to learn their tasks on the job, relying on the experience of other personnel for guidance; thus a good deal of variation in methodology may be expected. For this reason, estimative intelligence appears to differ from I&W intelligence, for which well-defined methods have been gathered in textbook form.

Estimators were concerned that, with the relatively high turnover of
personnel, each new estimator was required to develop his own methodology for
forming estimates. The lack of continuity is emphasized by the fact that
there is little overlap in personnel assignments; replacement personnel do
not usually have an opportunity to train with their predecessor on the job.

Estimators are constrained by the need to develop projections which will
be accepted by intelligence consumers. In part, they are cast in the role of
mediators among the various DOD and other intelligence producers, and their
estimates are not likely to be accepted if they deviate markedly from the
consensus of these sources. DE-1 plays an important role in judging the
plausibility of differing estimates.

7.2.   PROPOSED STATISTICAL METHODS

Several statistically-oriented methods have been proposed to assist in
making more accurate long-range projections. These have included such tech-
niques as the following:

1. Bayesian statistical analysis. Given appropriate prior probabili-
    ties and conditional probabilities, Bayesian analysis presents an
    effective method for determining the probabilities associated with
    stated outcomes. Although computational problems become severe as
    the number of elements in the evaluation becomes realistically
    large, various shortcut methods permit an approximation to the
    Bayesian solution for many problems.

2. Trend analysis. Econometrics has developed several techniques for detecting trends in time-series data and for projecting such trends into the future. It would be possible to develop extremely sensitive trend analysis routines through the use of statistical methods used for waveform analysis, e.g., the methods used for identifying individual radar signatures.

3. System dynamics. Computer-based methods for simulating large systems have been extensively investigated. Typically, these have modeled development of urban areas, industrial plants, or global ecology. There is strong interest in the application of system dynamics models to the production of long-range estimates.

4. Probability networks. Explicit probabilities are now being introduced into DE-1 reports. Through the use of such probabilities, it may be possible to obtain aggregate scores or probabilities for various hypotheses concerning weapon system projections.

Other statistical methods, such as correlation and regression, linear programming, the Delphi method, and PERT charts have been suggested for possible application to the production of DE-1 projections.

Statistical methods have not, however, been widely used in DE-1 estimates. The inclusion of the Kent charts as general guides to the probabilities associated with verbal statements of likelihood, and the use of some explicit

probability estimates in parentheses in reports, are the only apparent appli-
cations of mathematical probabilities.

An attempt was made to develop a model of the estimative methods actually
in use at DE-1, based on reports of methods used by estimators. Such a model
takes into account the start-up time required for design and initial produc-
tion of a new weapon system, productive capacities for the system, deployment
patterns, estimated useful lifetime of the system, and withdrawal or retire-
ment of the system. It was found that estimators tend to use such models as
guides in the development of projections for each weapon system.

In addition to mathematically quantifiable models, estimators also
rely on their knowledge of the subject matter. For example, personnel working
with Soviet projections learn that the tradition of five-year plans in the
USSR means that weapon development generally occurs in five-year blocks of
time. Again, it has been found that the Soviets tend to retain some equipment
beyond the point at which it would be regarded as obsolete by Americans; this
leads to a rule of thumb for estimating the time of phase-out of Soviet
equipment.

Occasionally, common-sense approaches are found to be inappropriate. For
example, it was assumed that Chinese forces would be expanded to include all-
weather aircraft; but the Chinese apparently did not feel the need for such
equipment.

In general, estimators appear to rely on a large number of informal models and rules to assist in the development of projections. At the start of the project, estimators were described as "non-statistically oriented"; this description appears to have been correct.

7.3.   ANNOTATIONS AND INTERPRETATIONS

Estimators were concerned to provide annotations for the proper justification and interpretation of projections. For some users, annotations may be more valuable than the numerical quantities that make up the DIPP tables.

In addition to the annotations, it was suggested that permanent files be maintained for the use of DE-1 personnel, containing interpretive material on the methods used in arriving at particular estimates. The files would contain a full statement of the assumptions, calculations, and background data that entered into each projection.

Such permanent files would serve three purposes:

1. With frequent turnover in DE-1 personnel, they would provide a source of information concerning the detailed methods used in production of DE-1 projections which would assist in training new personnel.

2. When questions concerning particular projections arise, it would be possible to determine sources of potential error in the assumptions

which entered into the projections. Since the person responsible
for a projection is not always available for questioning, the files
would provide the documentation needed for locating sources of
error.

3.   It may be expected that the quality of projections would be improved
      if estimators were required to provide detailed justification for
      them.

There is a great deal of information which is not in DIPPOLS, such as
deployment rate, the rate of change, the retirement rate, and the estimates of
ratios among weapon systems. Information such as this provides a context for
the figures which appear in the DIPP tables, and estimators expressed concern
that the figures should not be taken out of their context. A given projection
may be correct in saying, e.g., that a particular aircraft will be phased out,
even though the date of phase-out is somewhat inaccurate. Obviously, any
"scoring rule" provided with TEAMS should be used as a guide to locate sources
of trouble, rather than as a blind instrument which rates a projection as
"right" or "wrong" without considering the additional information provided in
the annotations.

7.4.   PROBLEM AREAS IN ESTIMATIVE INTELLIGENCE

In the remarks that follow, an attempt is made to place estimative
intelligence in a general context that includes basic, current, and I&W
intelligence functions, and to locate problems that appear to affect the

accuracy of long-range estimative intelligence. The character of long-range intelligence is such as to affect the nature of the error measurements that will be reported by TEAMS.

Traditionally, intelligence analysis has been faced with the task of inferring two unknowns: the capabilities and the intentions of a potential adversary. These two factors serve to determine the degree and nature of a potential threat, to which an appropriate response may be made. If the intelligence estimate is correct and timely, then the response will serve to reduce or eliminate the threat.

The dynamic character of this situation - in which intelligence serves to guide the response of operational forces - gives rise to the so-called "paradox of intelligence": "If intelligence is correct, then it will be proved wrong by events." That is, if an intelligence estimate is correct, it will serve to guide actions which reduce or eliminate the projected threat. Thus, intelligence cannot always be judged as "right" or "wrong," since its ultimate goal is to eliminate the threat which it predicts.

A version of this quasi-paradox affects long-range estimative intelligence. Long-range projections of the force levels of a potential adversary should have an effect on the development of U.S. force levels; these in turn will influence the actions of the potential adversary, with a resulting change in actual force levels. Thus, a "correct" estimate serves to falsify itself.

Long-range estimative intelligence is subject to two additional sources of apparent error, which do not affect short-range intelligence.

The first of these is the possibility of basic changes in attitudes of a potential adversary. Such a change may have occurred in the early 1960's, following the Cuban missile crisis. As a result of the U.S. confrontation, Soviet leadership may have made a change in policy, which led in turn to a much heavier reliance on ICBM emplacements. Since the change in Soviet policy had not been foreseen by the U.S., numbers of ICBM's were seriously underestimated for a time. Estimates, which may have been correct at the time they were made (in terms of Soviet intentions at that time), were proved to be incorrect when Soviet policy changed.

More generally, it may be said that long-range estimates are intrinsically subject to error, since the attitude of a potential adversary (and of one's own nation and its allies) can change in ways which are unpredictable in principle.

A second source of error which is unique to long-range estimative intelligence is the inherent unpredictability of technological development. An example of such development would be that of laser weaponry during the very recent past. Another development which would have been difficult to predict ten years ago has been the rapid miniaturization of computer components. The present role of computers in the U.S. would have been very difficult to predict in 1952, and, in fact, it was not predicted. As DE-1 extends the horizon

of its projections to include the next 25 years, this problem will grow in importance.

Thus, in addition to the sources of uncertainty which are endemic to all forms of intelligence, long-range estimates are subject to uncertainty in determining future attitudes and future capabilities of potential adversaries. While TEAMS is intended to detect errors in past projections, the meaning of "error" must be interpreted against the intrinsic sources of error in the task of estimative intelligence.

7.5.   COMMUNICATION OF UNCERTAINTY

A partial response to the problems of predicting future intentions and capabilities of a potential adversary has been the use of alternative sets of projections. One set of projections may be applied if it is assumed that an adversary will continue to abide by arms-limitation agreements, while another set of projections must be applied if the opposite assumption is made. Such an approach communicates to the user a sense of the sources of uncertainty concerning future weapons levels. However, it is hardly possible to antici-pate all future negotiations, agreements, and potential violations into which the nations of the world might enter over the next ten years.

More generally, a need exists for better methods for determining and communicating the uncertainty of long-range estimates. Estimators were particularly concerned with the development of better methods for indicating the degree of uncertainty in projections.

A specific example may be helpful. If we say that there is a 0.75 chance that the actual figure will lie between a given High and Low estimate, we are only partially communicating the uncertainty of this figure; information has been discarded. The uncertainty may reflect the probability attached to the overall force level, the probability of a particular mix of forces, the probability of deployment by a certain date, the probability of replacement by another system, the need to respond to U.S. actions, and so on. Although these probabilities can be combined to give a single number, such as 0.75, the user should also be aware of the reasoning that has entered into the determination of the projection. In short, the use of a single probability figure may be an extremely crude way of representing the analyst's actual thought processes.

An additional source of uncertainty was mentioned during the interviews but could not adequately be assessed. This was the degree of error which might be present in current order of battle (OB) data. Although estimators were aware that information concerning specific weapon systems in certain countries was doubtful or inaccurate, there did not seem to be a clear way of incorporating knowledge about such inaccuracies into the projections. It should be noted, however, that "errors" in some projections could be due in part to errors in the OB data; such errors in the OB would affect projections which were based on OB figures, as well as evaluations of prior projections which were compared with current OB figures.

To take a hypothetical example, if there were a sharp increase in ICBM figures from 1972 to 1973, this might be due to an improvement in collection

techniques, rather than to an increase in actual numbers of ICBMs. In addi-

tion, projections which were based on the 1972 figures would be in error when

evaluated against the 1973 data. Finally, projections made in 1970 might be

judged as "errors" when they were compared with the 1972 figures, but could be

judged as correct when compared with 1973 information. In short, errors in

the OB can have a pervasive effect upon judgments of the accuracy of projec-

tions.

SECTION 8

RECOMMENDATIONS FOR FUTURE DEVELOPMENT

The TEAMS project has provided designs for a system which is intended
to assist DIA-DE-1 estimators and managers in identifying and quantifying
past errors, and for comparing past estimates, to provide a record of their
successes and a means for reducing their errors. It is hoped that TEAMS
will prove to be an effective tool in the development of techniques for
estimative intelligence.

TEAMS project personnel have become deeply involved in the study of
tasks performed by DE-1. The importance of these tasks, which have a sig-
nificant role in determining this country's military posture, together with
the difficulty inherent in providing projections of future developments,
have given the project an unusual degree of interest.

In this section, which forms a conclusion to the Final Report, we will
review the additional work that will be required to complete TEAMS as an
operational data management system. Coordination and consolidation with the
existing DIPPOLS system are also proposed. In addition, several areas of
methodology have been reviewed during the project, and these are described
here as potential research efforts.

8.1.   DIPPOLS REDESIGN

DIPPOLS is in daily use by DE-1 and other agencies. It differs from
TEAMS, in that TEAMS is intended for the exclusive use of DE-1 in preparing
reports and in self-evaluation. Nevertheless, DIPPOLS and TEAMS share much
of the same data, and DIPPOLS will act as the source of information for the
TEAMS data base.

Development of TEAMS designs has required a comprehensive review of
DIPPOLS, and has led to a number of suggestions for modification and redesign.
Briefly, these include the following:

1.   Data storage and access could be modified to develop a data base
     which more closely resembles the TEAMS data base. In particular,
     the hierarchical system of pointers used in TEAMS gives more rapid
     and effective access to the data.

2.   Since a substantial proportion of the DIPPOLS data base is included
     in the TEAMS data base, it might be desirable to consolidate the
     two data bases. A further step would be consolidation of the two
     systems, which would serve to eliminate duplicated functions.

3.   Other modifications of DIPPOLS could include improvements in the
     command language, modification of the general program design, and
     enhancement of DIPPOLS report generating functions.

8.2.   TEAMS IMPLEMENTATION

Implementation of the TEAMS design is the next step in TEAMS develop-
ment. This project will be the straightforward task of coding and testing
the routines documented in the Program Specifications which accompany this
Final Report.

To provide the basis for an estimate of time and cost for TEAMS imple-
mentation, Table 8-1 includes a summary of components of the system with the
approximate numbers of lines of FORTRAN code that will be required for final
implementation. Note that the TEAMS Pilot, which was used for experimental
testing of system designs and for demonstration, consists of 34 routines
comprising 1661 lines of code (LOC), and that the final TEAMS implementation
will require an estimated 5227 LOC.

8.3.   TEAMS EXTENSIONS

The TEAMS design may be expanded to include several additional facilities.
Among those of greatest interest are the following.

1. System use monitoring. This facility would provide statistics on
    use of TEAMS subsystems, to locate those which are using excessive
    computer time, to identify those which are unused or underused, to
    obtain records of error returns, etc. The purpose of the monitor
    would be to assist in improving TEAMS performance.

Table 8-1 Lines of Code Estimated for TEAMS Implementation

I.   TEAMS PILOT

        a)   Consists of 34 routines

        b)   or equivalently 1661 lines of code (LOC)

        LOC (1) = 1661

II.   ADDITIONS TO TSS (TIME-SHARING SYSTEM) ROUTINES

Routine	LOC
SELECV	100
MODIFY	350
SORT/RANK	360
PLOT	450
PRINT	100
LOCK/UNLOCK	10
BIAS/ERROR/UNCERT	163
COMBINE	251

Table 8-1 (Continued)

Routine	LOC
PARAMPD/J	204
SCORE	100

LOC (2) = 2088

III. BATCH PROGRAMS

Routine	LOC
LPRINT	20
PLOT	400
DURC	150
UPDATE	450
TIO	50

LOC (3) = 1070

IV. ADDITIONS TO EXISTING ROUTINES (I.E., THE PILOT)

Routine	LOC
ASSIGN (variable names)	56
KPARSE (full key PARSE)	112
DIPSEE (exclusions)	50

Table 8-1 (Continued)

Routine	LOC
Partial Projections	20
Extensive Error Messages	50
Additions to Chief	10
Lock Check in STOWN, ERASE	20
CHOMP (line continuation)	20
CLEAR (confirmation check)	10
MAP (bit map)	40
PEEK	10
DSPUAR (plot format print)	10

LOC (4) = 408

Total Lines of Code = 5227

2. Data Base Expansion. DIPPOLS and TEAMS contain only part of the data used by DE-1 in developing projections. A more comprehensive data management system for DE-1 would include such additional information as factory production rates, deployment rates, and other information. An important part of this development would be effective means for accessing the new data, comparing it with data presently contained in DIPPOLS, and outputting reports and other information in useful formats.

3. Arithmetic Functions. At the present time, DE-1 estimators use hand calculators for most of the arithmetic operations required for developing and modifying projections. Although such use is more convenient for them than the use of a computer terminal, which is located in another room, is not always available, and requires several minutes to initiate, nevertheless, as computer use increases, it will also be desirable to increase the range of functions which are available from the computer terminal. Such facilities would permit TEAMS users to perform a range of algebraic and arithmetic operations on TEAMS data directly at the terminal.

4. Subscripted Variables and Multi-Dimensional Data Structures. The availability of subscripted variables and N-dimensional data structures would enhance the command language and provide greater computational power, respectively.

5. Enhancements in Plotting and Table Generation. During TEAMS implementation and testing, it is likely that table generation and plotting facilities, which were not included in the TEAMS designs, will be found desirable. Users and potential users of the system will doubtless find many additional functions for potential inclusion. It would be desirable to collect such suggestions and to plan for expanded TEAMS facilities in response to them.

6. Use of DIAOLS Routines. As a part of the TEAMS effort, a review of analytic and statistical packages currently available on the DIA computer systems was conducted. This review did not indicate that such packages could be incorporated in the TEAMS design, given the requirement that TEAMS be directed to non-statistically oriented estimators and managers.

It is nevertheless likely that some DE-1 personnel may find it desirable to perform statistical studies on TEAMS data, such as regression analysis, factor analysis, analysis of variance, etc. Under these conditions, facilities might be made available for the transformation of TEAMS data into formats acceptable to the DIA analytic and statistical packages now available through DIAOLS.

8.4.   ADDITIONAL HARDWARE FACILITIES

The TEAMS design was developed for existing DIA computer facilities, together with existing terminals and line printers. Plotter facilities have been projected on the assumption that a standard incremental plotter would be made available. Other types of hardware facilities, specifically intended for TEAMS/DIPPOLS use, might include the following:

1. A dedicated computer system located at DE-1 offices, utilizing a minicomputer such as the AN/GYQ-21(V) analyst station, could provide much more convenient facilities for the production of required TEAMS and DIPPOLS outputs. With the decreasing costs of minicomputer systems and the heavy load of the DIA central facilities, it would be desirable to develop cost studies of such an implementation.

2. It would be desirable to review CRT display units as potential terminals for DE-1 use. With a visual graphic display, the user would be able to see proposed plots immediately, and to change parameters as required to produce the desired output. With hard-copy production facilities, the user would also be able to obtain graphic output directly from the display terminal.

1.0

1.1

1.25　　1.4　　1.6

4.5
4.0
3.6

2.8　　2.5

3.2　　2.2

3.6

4.0　　2.0

1.8
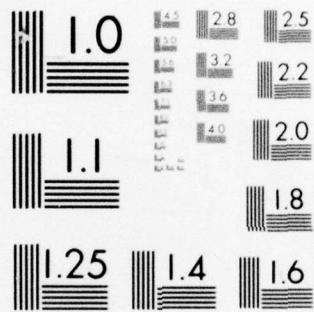
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

8.5. RESEARCH IN ESTIMATIVE INTELLIGENCE METHODOLOGY

In this subsection, several research projects are briefly outlined as examples of work that could assist in improving projection techniques.

8.5.1. Trend Analysis Research

A substantial proportion of estimative intelligence analysis is concerned with identification of trends, such as trends in deployments to selected areas, industrial production, composition of forces, utilization of new technologies, etc. Statistical trend analysis has traditionally been used for this purpose, to locate and quantify various components of the trend, such as secular trends, seasonal or weekly cycles, random noise, and significant departures from previous levels.

Mathematical models of trend analysis have generally been extremely simple; this has been particularly true of econometric models, where complexity enters only because of the large number of variables involved. Recent research in pattern recognition for waveform analysis has, however, provided more sensitive mathematical models for the measurement of trends.

The goal of this research would be the application of waveform analysis techniques to the analysis of trends of interest to estimative intelligence. TEAMS will include historical data concerning foreign weapon systems over a period of several years, which may be applicable to this study. Emphasis

during initial research would be on technology transfer, to permit the use
of tested methods of waveform analysis in the area of estimative intelligence.

8.5.2.   Quantification and Communication of Uncertainty

Like other major types of intelligence, estimative intelligence deals
with information which is known to be uncertain. Sources of uncertainty
include resolution levels of collectors, inaccessibility of required infor-
mation, conscious deception by an adversary, concealment or changes of an
adversary's intentions, and so on.

The Sherman Kent chart was developed as a tool for communicating the
degree of uncertainty to be ascribed to information in a firm, quantifiable
way, by translating such natural-language terms as "almost certain," "prob-
able," "doubtful," etc., into numerical probabilities. Unfortunately, such
probabilities frequently give an unwarranted sense of precision to the
estimates. Many studies have demonstrated that there is a non-linear rela-
tion between subjective probabilities and the actual probability of an
event.

Recent academic research has developed effective statistical techniques
for dealing with imprecise definitions. Such techniques replace the all-or-
nothing techniques of traditional set theory with methods for dealing with
partial set memberships. Development of these techniques has provided a
tool for communicating the uncertainty to be ascribed to intelligence estimates

and judgments, and for combining these judgments in more inclusive estimates of the probability of future force levels.

8.5.3. Development of Training Materials

Training of DE-1 personnel has tended to take two forms: (1) On the one hand, personnel have often had to learn on the job, from their associates and supervisors. Because of the frequent turnover of military personnel, such training represents a continuing requirement; military assignments often are such that there is no opportunity for a replacement person to train with his predecessor. (2) On the other hand, existing training materials include statistical techniques which non-statistically oriented personnel have found difficult to apply to their work.

Existing texts provide training materials for other types of intelligence analysis; an example would be Indications and Warning Intelligence (U), developed for the Defense Intelligence School (Jan. 1975, 3 Vols.). Through a review of actual and potential methods in use in estimative intelligence, it should be possible to develop appropriate training materials of a similar depth. The purpose of this effort would be not only to assist in training of new personnel, but to assist in providing some consistency in the terminology and methods used by estimators and managers.

8.5.4.   Systems Simulation for Estimative Intelligence

In developments such as system dynamics, computer-assisted simulations of large systems have been studied extensively. Simulations of industries, urban areas, and global ecological systems have provided a basis for studies of the future of these systems, and of the effects of changes in the parameters affecting them.

While such studies are not likely to replace human judgments in estimative intelligence, they can nevertheless assist the human estimator in providing a rapid test of various hypotheses concerning future developments. It would be possible, for example, to determine the effects of declines in availability of petroleum and other resources, changes in industrial productivity, and many other variables upon estimates of weapon system levels.

It would be desirable to investigate the possibility of making such simulation systems available to the DE-1 estimator to assist in production and evaluation of projections.

APPENDIX A

TEAMS TERMINOLOGY

PREDICTION (PD) — A prediction is an estimate of a future force level, for a particular weapon, for a specific year.

PROJECTION (P) — A projection is an annually issued, chronologically ordered sequence of eleven* predictions plus OB values for the prior four years. There are three types of projections.

    a. The high projection (H) which consists of the highest predicted values (i.e. force levels) expected for each of the eleven prediction years plus the prior OB values.

    b. The low projection (L) which consists of the lowest predicted values (i.e. force levels) expected for each of the prediction years plus the prior OB values.

    c. The best projection (B) which consists of the most probable predictions (i.e. force levels) for each of the prediction years plus the prior OB values.

—————————

* TEAMS actually allows for a variable number of predictions and OB values to be contained in a projection.

PREDICTION YEAR   -   A prediction year is the year <u>for which</u> a prediction
(PDY)   was made.

PROJECTION YEAR   -   A projection year is the year <u>in which</u> a projection
(PJY)   was made.

ORDER OF BATTLE   -   Each <u>issue</u> of a projection provides (in addition to
(OB)   predictions) a current best estimate of actual
                          force levels for the 4 previous years (i.e. the 4
                          years prior to the PJY). The current best estimates
                          of force levels for all weapon systems for past
                          years are called the OB values for those years.

ORDER-OF-BATTLE   -   An Order-of-Battle year is the year associated with
YEAR (OBY)   the OB.

BEST OB (BOB)   -   Best OB is the latest OB available for any OBY
                          prior to the current projection year.

PARTIAL   -   A partial projection refers to a portion of an
PROJECTION (PP)   entire projection.

A-2